DOE transforms HPC

Pacific Northwest NATIONAL LABORATORY

PNNL's Open Community Runtime

High Performance Data Analytics

Whole Program Adaptive Error Detection and Mitigation (AEDAM)

Advanced Network Architectures: Data Vortex

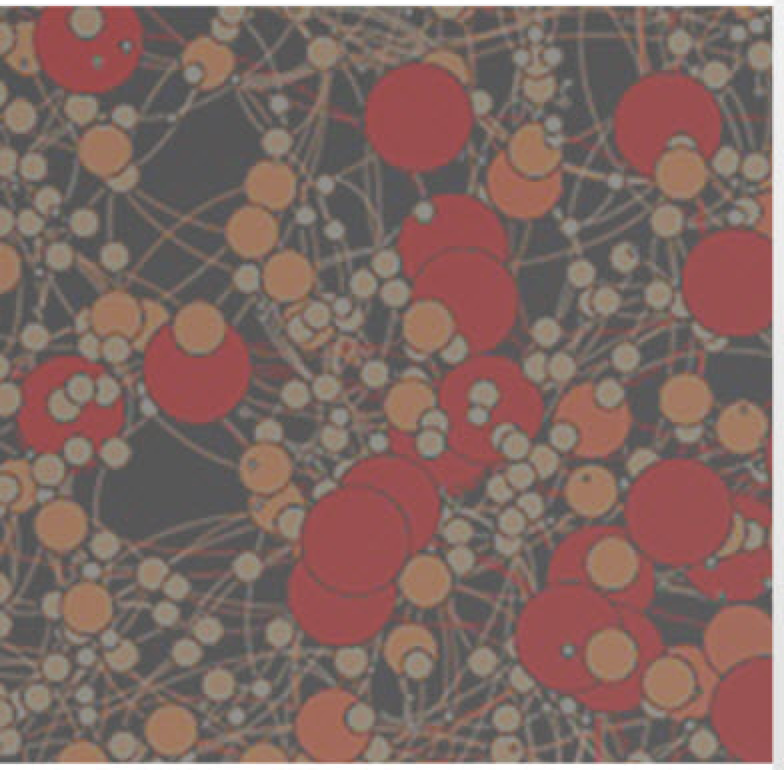
Integrated End-to-End Performance Prediction and Diagnosis for Extreme Scientific Workflows (IPPD)

BACK TO MAP

Modeling HPC Workload Behavior

Assessing the Impact of Silicon Photonics for Graph Applications

Pacific Northwest NATIONAL LABORATORY



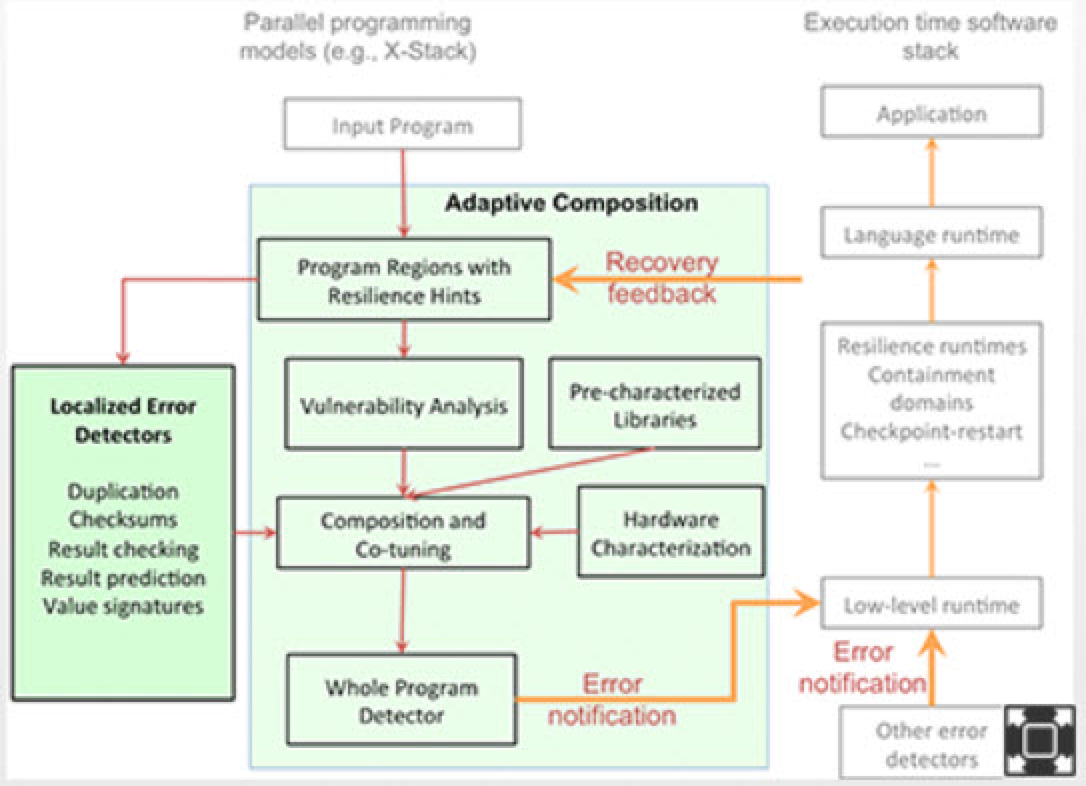
Whole Program Adaptive Error Detection and Mitigation (AEDAM)

Errors in hardware that escape detection can lead to silent data corruption. Detecting these types of errors is an important first step toward fault-tolerant program execution.

PNNL, in partnership with several leading academic research institutions, is designing a comprehensive approach to error detection and mitigation for scientific applications. The approach combines configurable error detectors, a unified reliability specification, and whole-program detector composition.

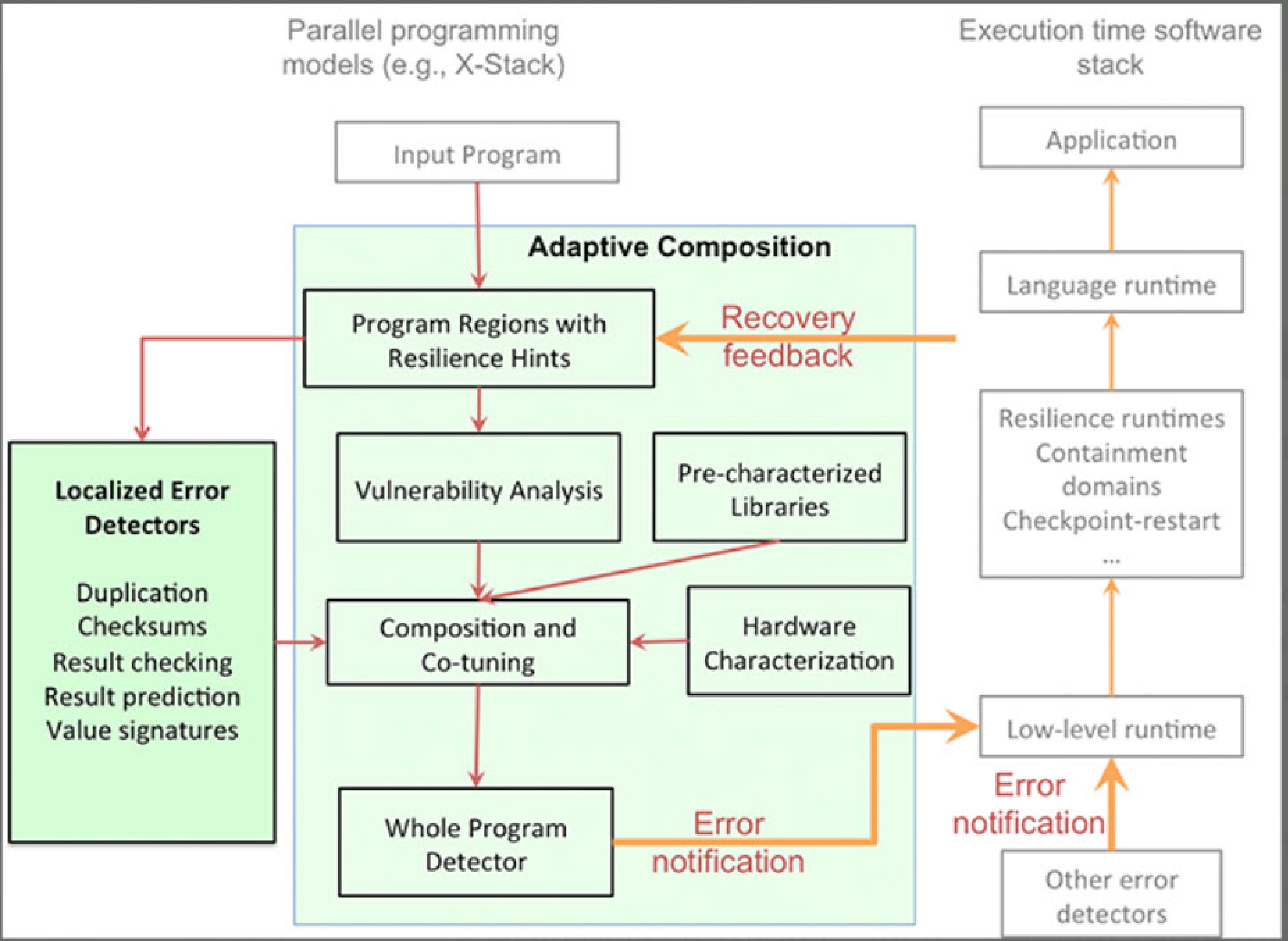
Specifically, AEDAM involves:

- Designing efficient and effective detectors for different portions of the execution state in scientific applications
- Characterizing and composing individual detectors to design an effective full-application solution
- Evaluating and understanding detector compositions
- Evaluating the impact of fault rates on the balance between hardware and software detection strategies.



Pacific Northwest NATIONAL LABORATORY · MIT · THE UNIVERSITY OF TEXAS AT AUSTIN · UNIVERSITY OF MICHIGAN · THE OHIO STATE UNIVERSITY · THE UNIVERSITY OF UTAH

FUNDING & CREDITS



Pacifi
N

rror Detecti

o silent data corruption.
xecution.

research institutions, is
ions. The approach com
r composition.

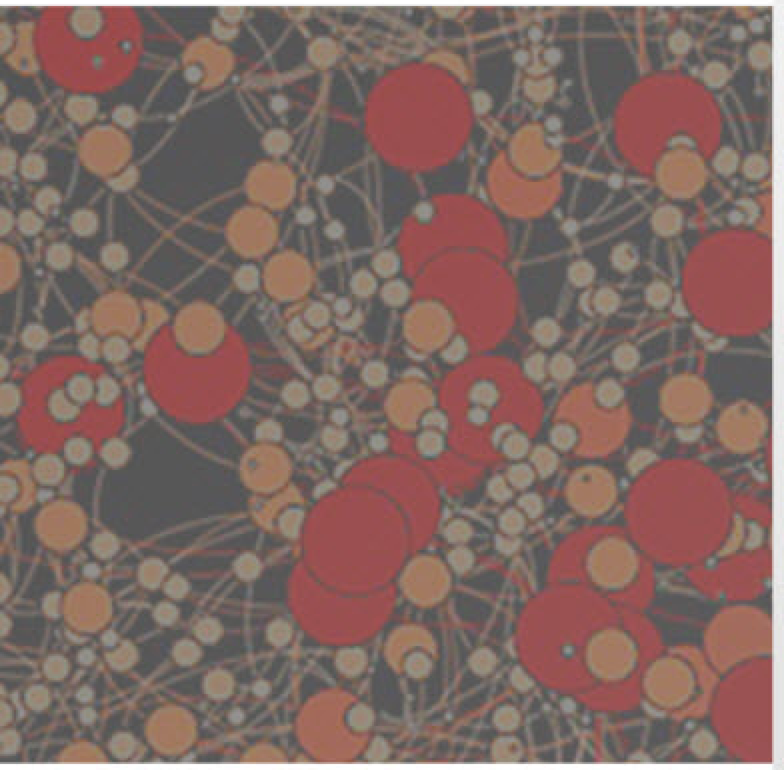
ferent portions of the exe
to design an effective f
ns
between hardware and software detection strategies.

BACK TO MAP

Funding for AEDAM is provided by the U.S. Department of Energy's Office of Advanced Scientific Computing Research. PNNL is operated by Battelle for DOE under Contract DE-AC05-76RL01830. (PNNL-SA-113784)
Contact: Sriram Krishnamoorthy, sriram@pnnl.gov

FUNDING & CREDITS

Pacific Northwest NATIONAL LABORATORY



BACK TO MAP

PNNL's Open Community Runtime

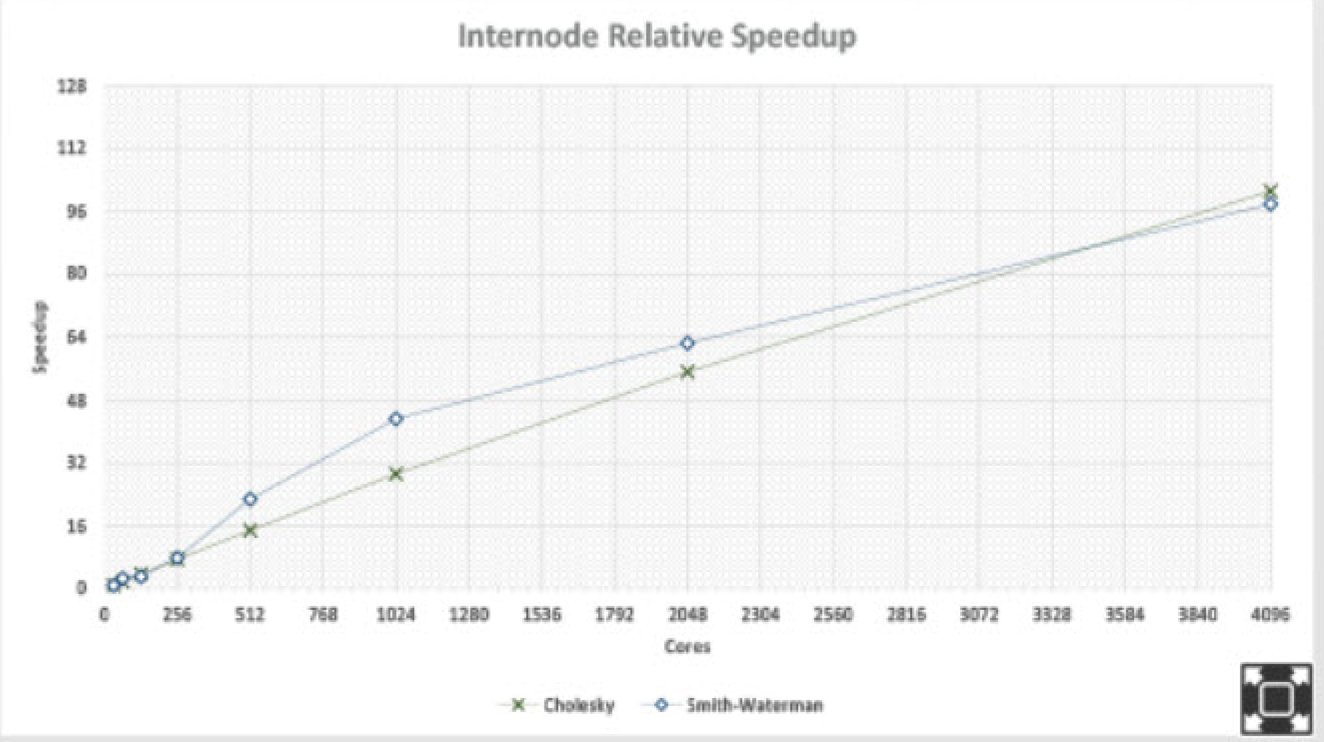
P-OCR: An asynchronous event-driven runtime that finally flies

With the constraints of energy, data movement, and resiliency, future many-core architectures will impose new challenges on applications when scaling across hundreds of cores within a node and even greater challenges when scaling across thousands of these nodes.

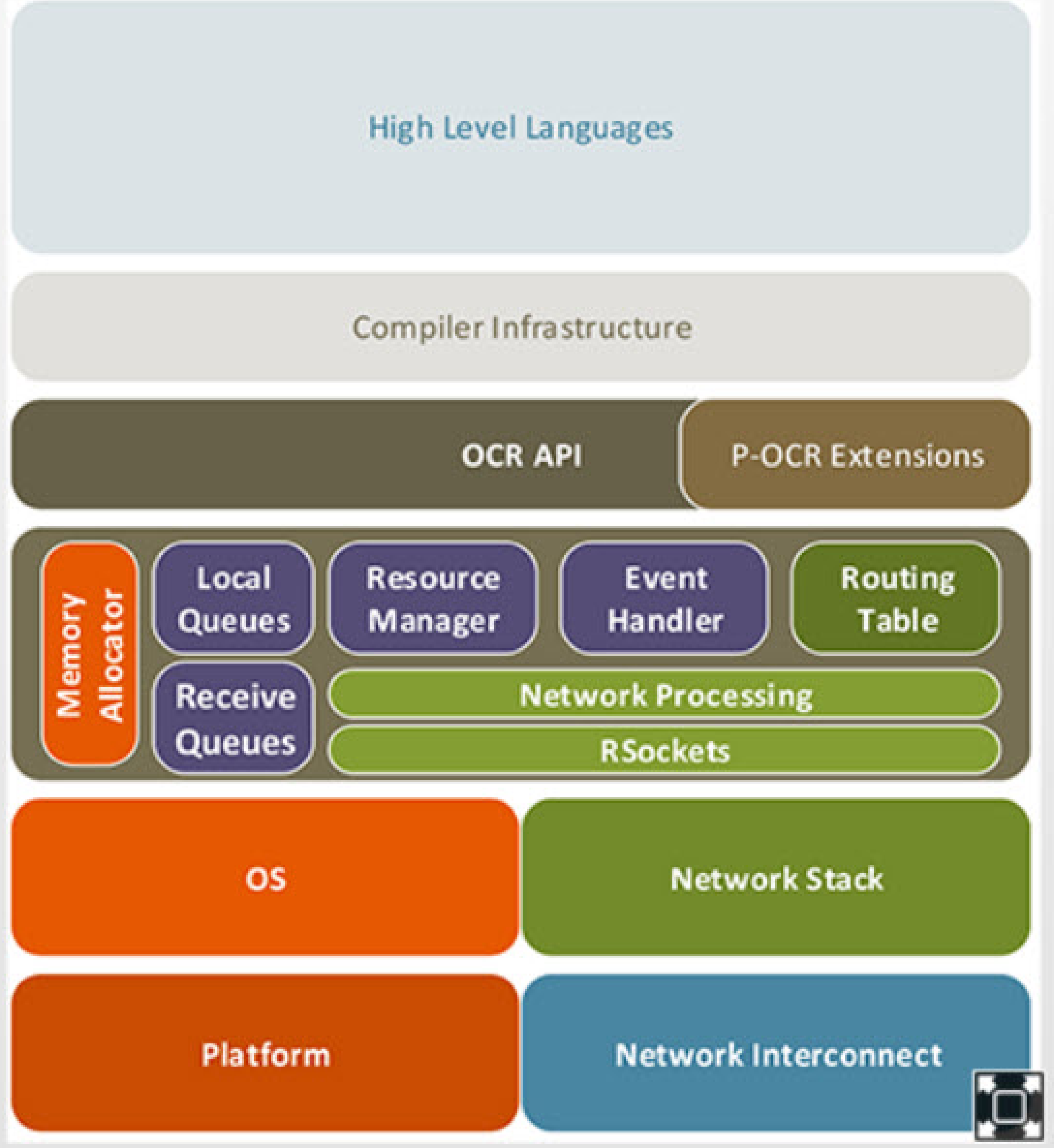
Pacific Northwest National Laboratory's Open Community Runtime, or P-OCR, is a framework that implements the OCR standard across current distributed systems, exploring how fine-grained, event-driven tasks; data movement; and dynamic resource adaption provide scalability for regular and irregular applications while managing those constraints.

With P-OCR on a general-purpose cluster, we have demonstrated up to 1592x speedup for a parallel-tiled Cholesky kernel and 2380x speedup for a parallel-tiled Smith-Waterman kernel using 4096 cores.

P-OCR scales within a node using shared memory* and multiple threads for both irregular and regular applications (*every two cores share a floating point unit).

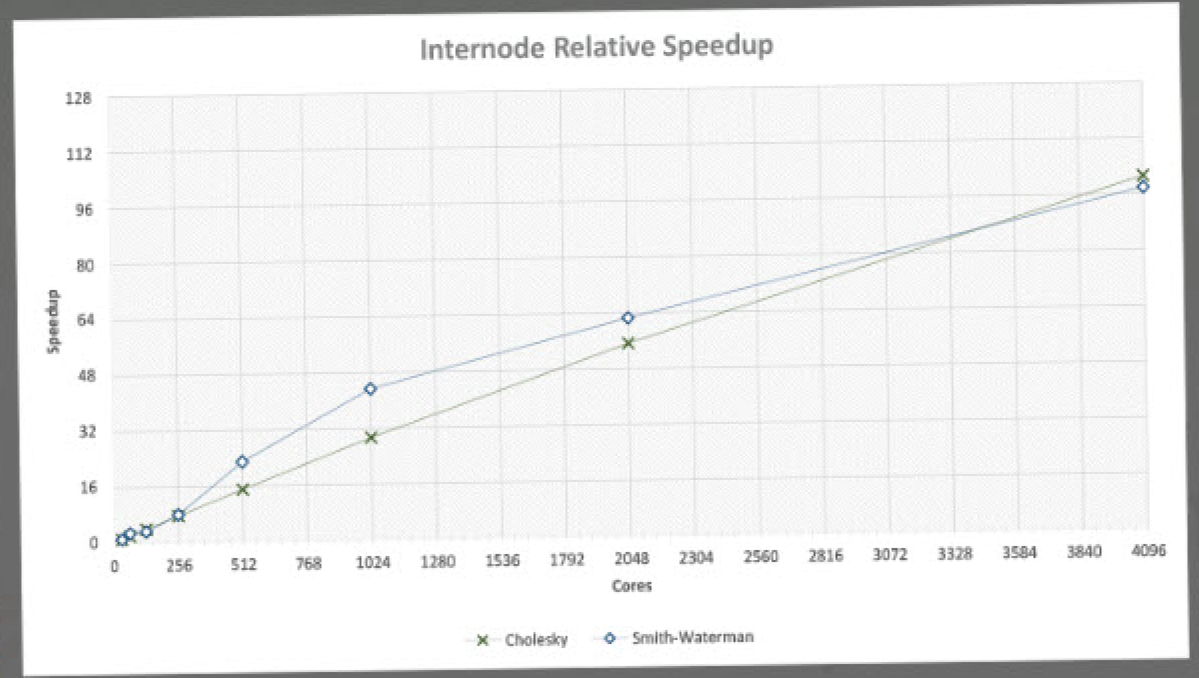


P-OCR scales across nodes using Rsockets and InfiniBand. It can achieve 100x speedup on 128 nodes over one node for irregular applications.



P-OCR is a tightly coupled multithreaded runtime that combines both distributed and shared memory features to achieve scalability within or across nodes.

FUNDING & CREDITS



P-OCR scales across nodes using Rsockets and InfiniBand. It can achieve 100x speedup on 128 nodes over one node for irregular applications.

Pacific Northwest NATIONAL LAB

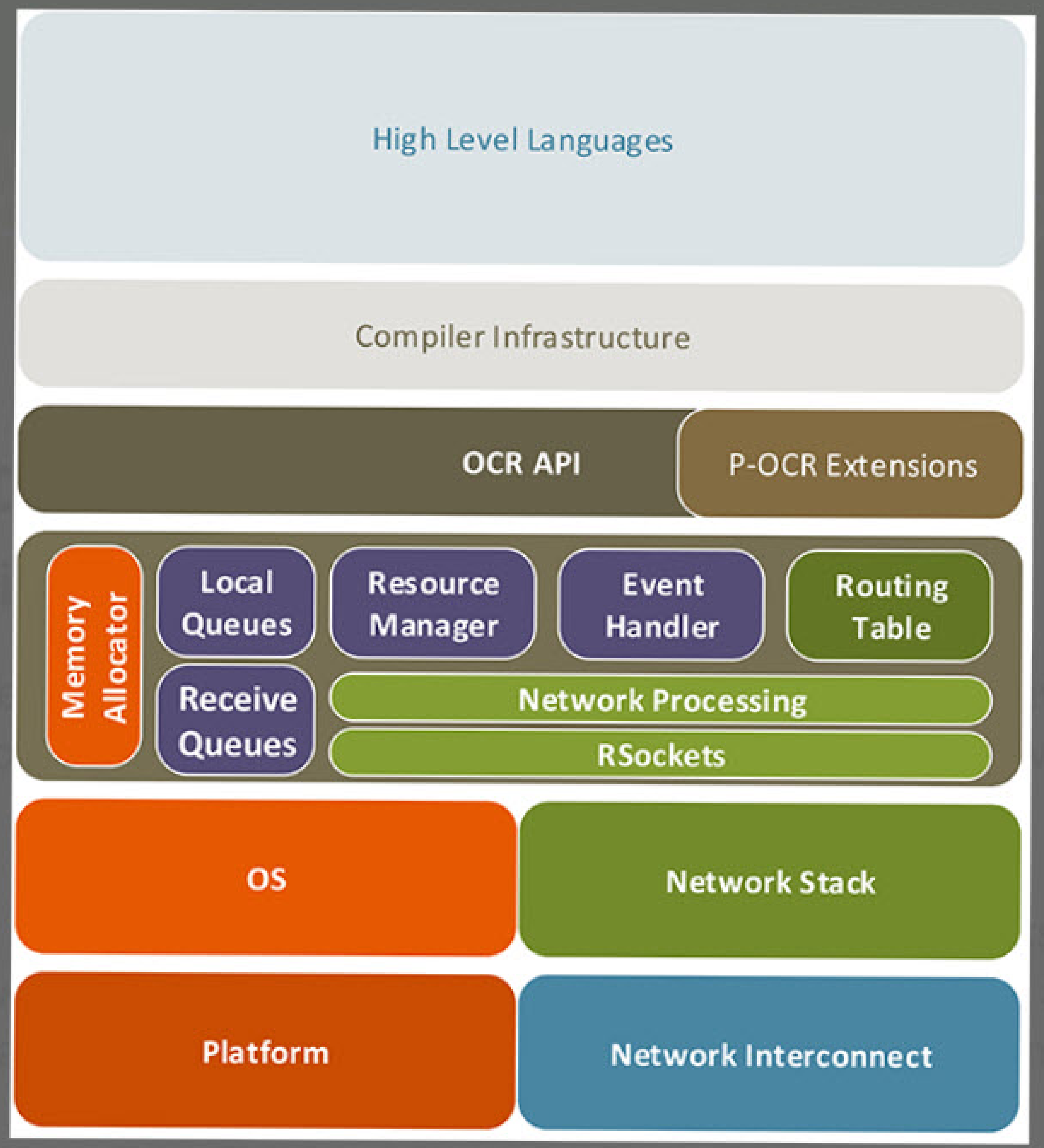
munity Runt

event-driven runt

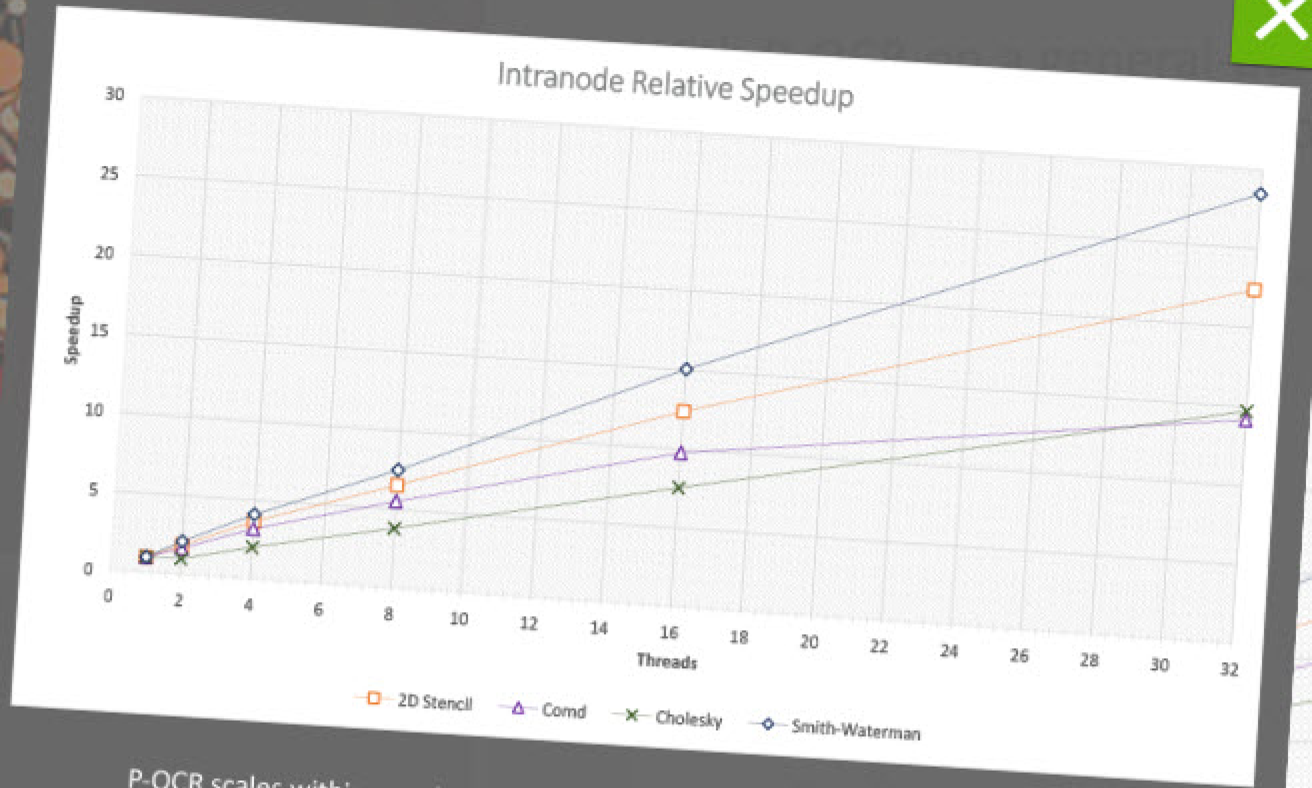
ta movement, and resili ... challenges on applications when scaling across hundreds ... scaling across thousands of these nodes.

Pacific Northwest National Laboratory's Open Community ... the OCR standard across current distributed systems, exploring ... dynamic resource adaption provide scalability for ... straints.

... ose cluster, we have demonstrated up to 1592x speedup for a parallel-tiled Cholesky ... r a parallel-tiled Smith-Waterm...

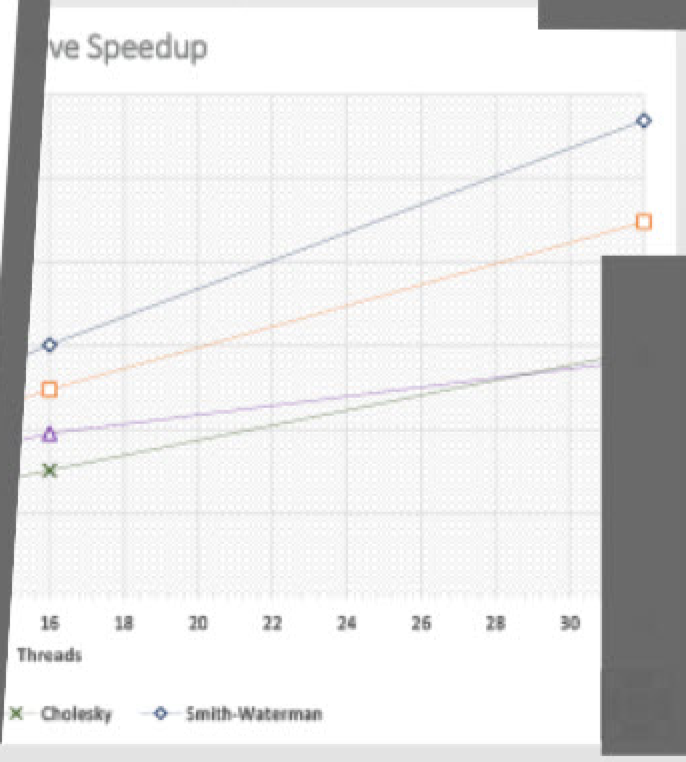


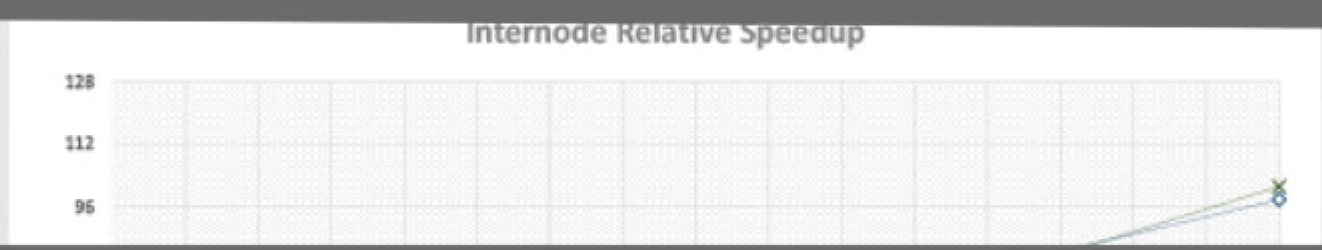
P-OCR is a tightly coupled multithreaded runtime that combines both distributed and shared memory features to achieve scalability within or across nodes.



P-OCR scales within a node using shared memory* and multiple threads for both irregular and regular applications (*every two cores share a floating point unit).

BACK TO MAP

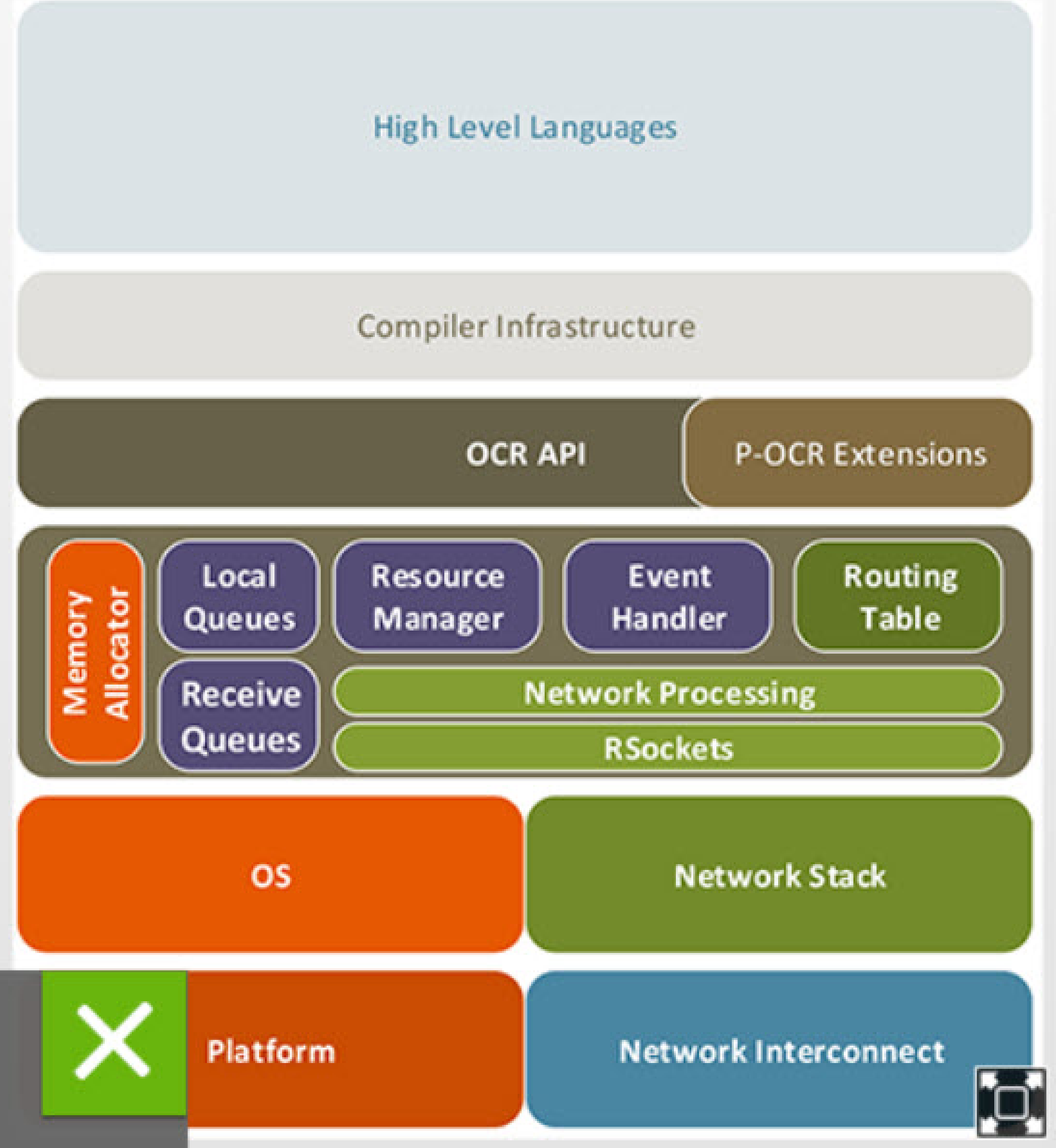




Funding for P-OCR is provided by the U.S. Department of Energy's Office of Advanced Scientific Computing Research. PNNL is operated by Battelle for DOE under Contract DE-AC05-76RL01830. (PNNL-SA-113783)
Contact: Andres Marquez, andres.marquez@pnnl.gov

P-OCR scales within a node using shared memory* and multiple threads for both irregular and regular applications (*every two cores share a floating point unit).

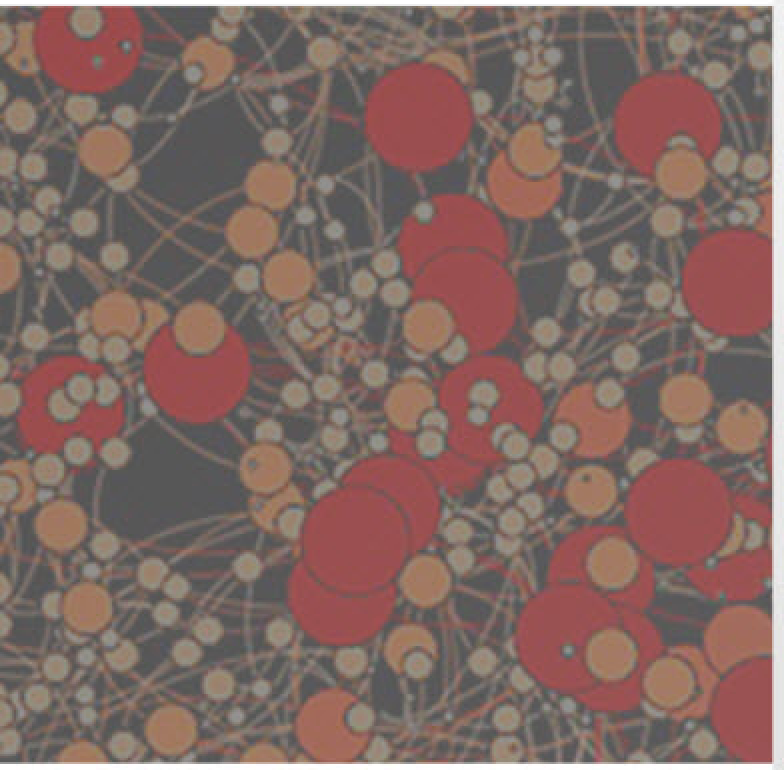
P-OCR scales across nodes using Rsockets and InfiniBand. It can achieve 100x speedup on 128 nodes over one node for irregular applications.



tightly coupled multithreaded runtime that both distributed and shared memory features to alability within or across nodes.

FUNDING & CREDITS

Pacific Northwest NATIONAL LABORATORY



BACK TO MAP

Advanced Network Architectures: Data Vortex

Pacific Northwest National Laboratory is the first site to install a Data Vortex system.

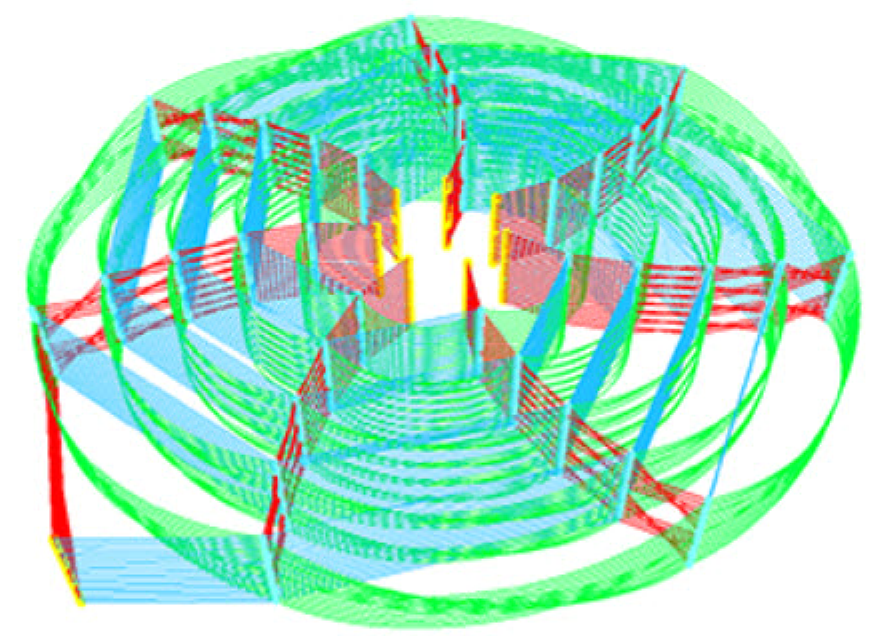
The DV205 system, dubbed “PEPSY,” is a 32-node cluster. Each node includes two Quad-Core Intel Xeon E5-2623 v3 at 3.0 GHz and 160 GB of RAM. The nodes are interconnected through both a conventional InfiniBand network and a custom network designed by Data Vortex. The custom design is a high-radix network with high bandwidth and low latency that supports very fine-grained transactions virtually congestion-free.

The design includes a Vortex Network Interface that implements a static random-access memory (SRAM)-based local memory and a custom network switch.

The Data Vortex Network may afford ways to build more balanced and efficient networks with applications and algorithms that present fine-grained data accesses, including Fourier analysis, sparse linear algebra, partial differential equations, and analytics on large data sets.

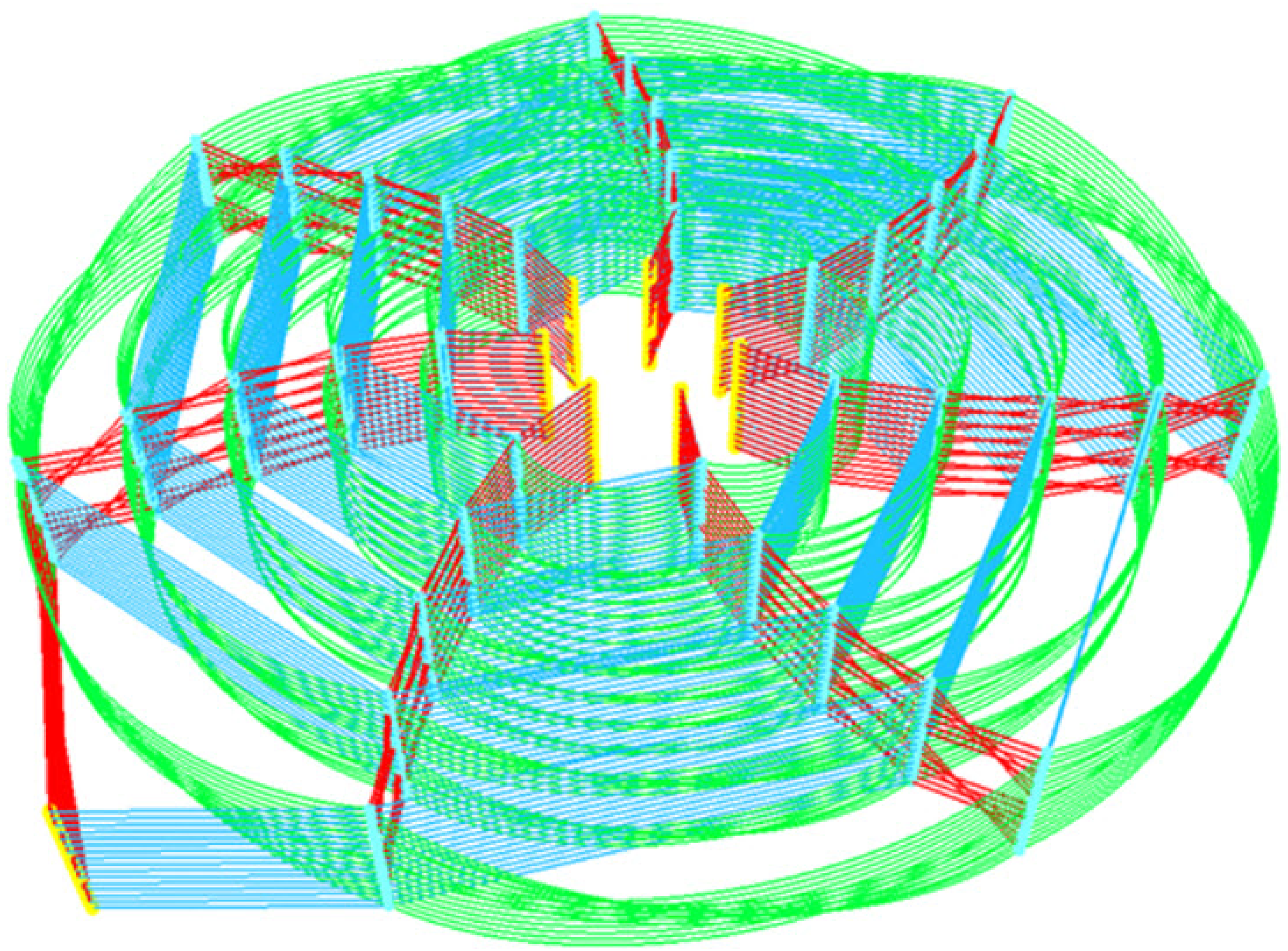
Dr. Coke Reed, chairman and founder of Texas-based Data Vortex Technologies, invented the Data Vortex Switch design from problem 110 entered by Stanislaw Ulam in the Scottish Book. Reed jointly solved the problem, concerning the existence of a fixed point in whirlpool-like flows, with Dr. Krystyna Kuperberg in 1976.

The Data Vortex Switch Topology



FUNDING & CREDITS

The Data Vortex Switch Topology



…itectures: D…

…site to install a Data…

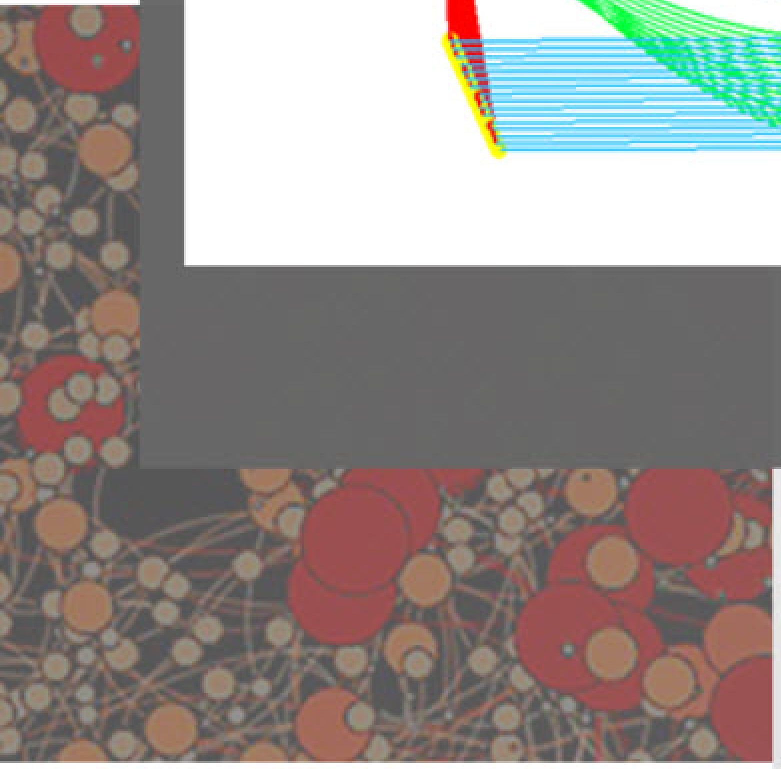
…e cluster. Each node i…
…connected through b…
…stom design is a high-…
…ons virtually congesti…

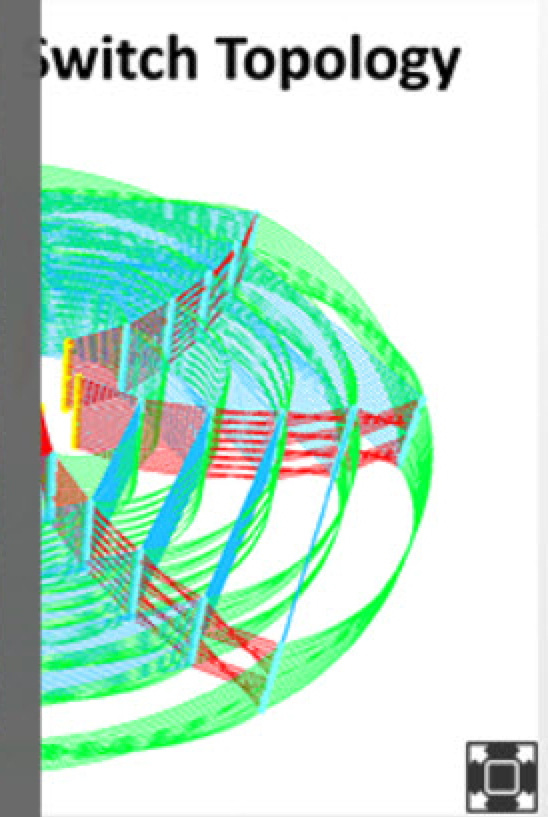
…hat implements a sta…

The Data Vortex Network may afford ways to build more balanced and … algorithms that present fine-grained data accesses, including Fourier analysis, sparse linear algebra, partial differential equations, and analytics on large data sets.

Dr. Coke Reed, chairman and founder of Texas-based Data Vortex Technologies, invented the Data Vortex Switch design from problem 110 entered by Stanislaw Ulam in the Scottish Book. Reed jointly solved the problem, concerning the existence of a fixed point in whirlpool-like flows, with Dr. Krystyna Kuperberg in 1976.

Pac…



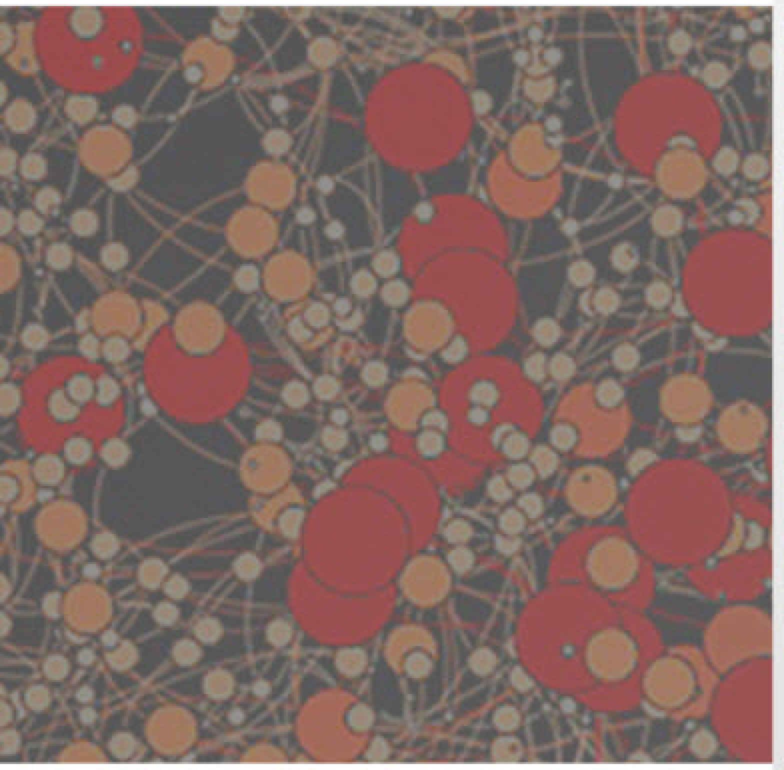


This technology benefits work supported by the High Performance Data Analytics Program, a collaboration between PNNL, Mississippi State University, University of Washington, and Georgia Institute of Technology. PNNL is operated by Battelle for the U.S. Department of Energy under Contract DE-AC05-76RL01830. (PNNL-SA-113909)
Contact: David Haglin, david.haglin@pnnl.gov

BACK TO MAP

FUNDING & CREDITS

Pacific Northwest
NATIONAL LABORATORY



BACK TO
MAP

Modeling HPC Workload Behavior

Performance Analysis of Large-Scale Systems and Applications

Application modeling captures the behavior of full-scale workloads and the interactions between these applications and the underlying hardware platform. PNNL's models allow researchers to navigate complex application and system design spaces in a rapid fashion, focusing efforts on design options that will yield the greatest benefit. The Modeling Workflow (above) can be broken into three activities:

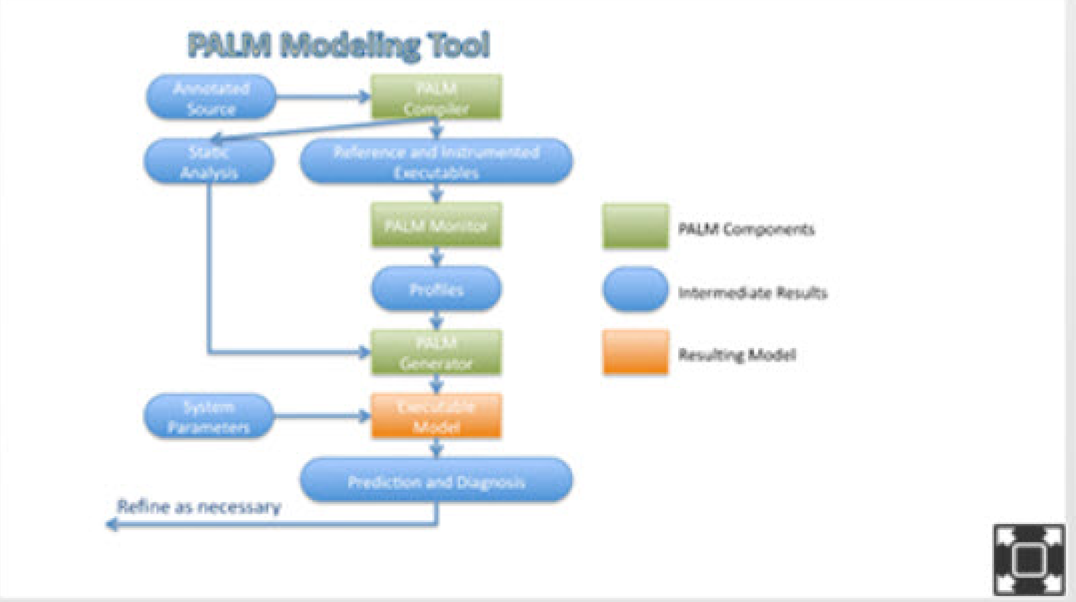
Workload and architectural specifications: targeting interesting point design studies.

Behavioral analysis and mapping: characterizing key application behaviors (e.g., computation, memory access, communication) and mapping tasks onto resources.

Prediction and validation: evaluation of models to derive predictions and validation of predictions against available measurements or simulations.

PNNL currently is advancing the state-of-the-art in modeling by developing modeling tools, applying modeling to distributed workflows, evaluating new parallel execution models, and other efforts related to U.S. Department of Energy and Department of Defense-funded programs.

Modeling requires expertise and labor that may not be available to application and system researchers. The PALM tool aims to simplify the modeling process, using an annotation language expresses insight and guides modeling. The model and application can be developed in tandem. Simultaneously, routine tasks may be automated. The resulting models are executable, reproducible, and distributable.

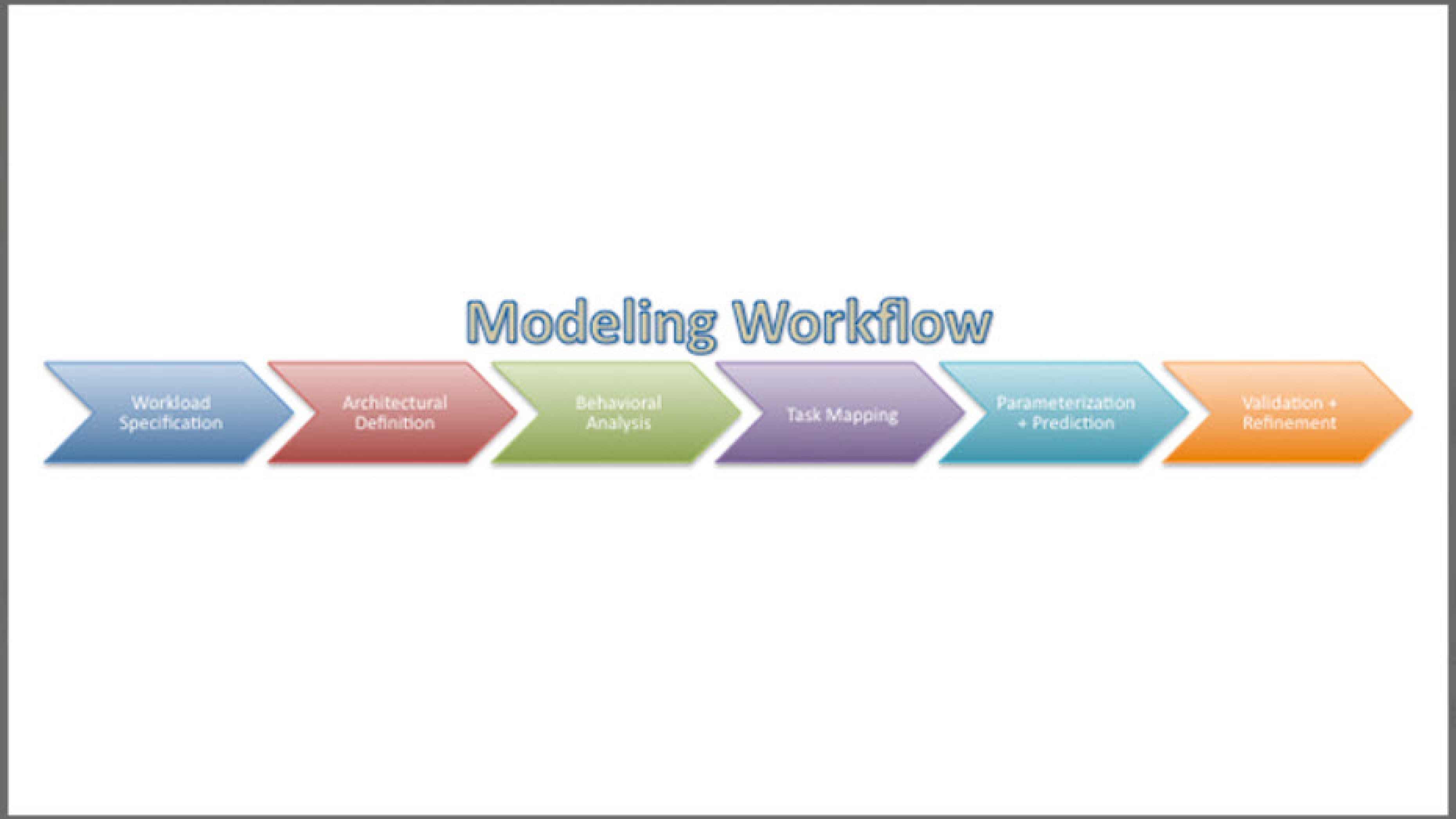


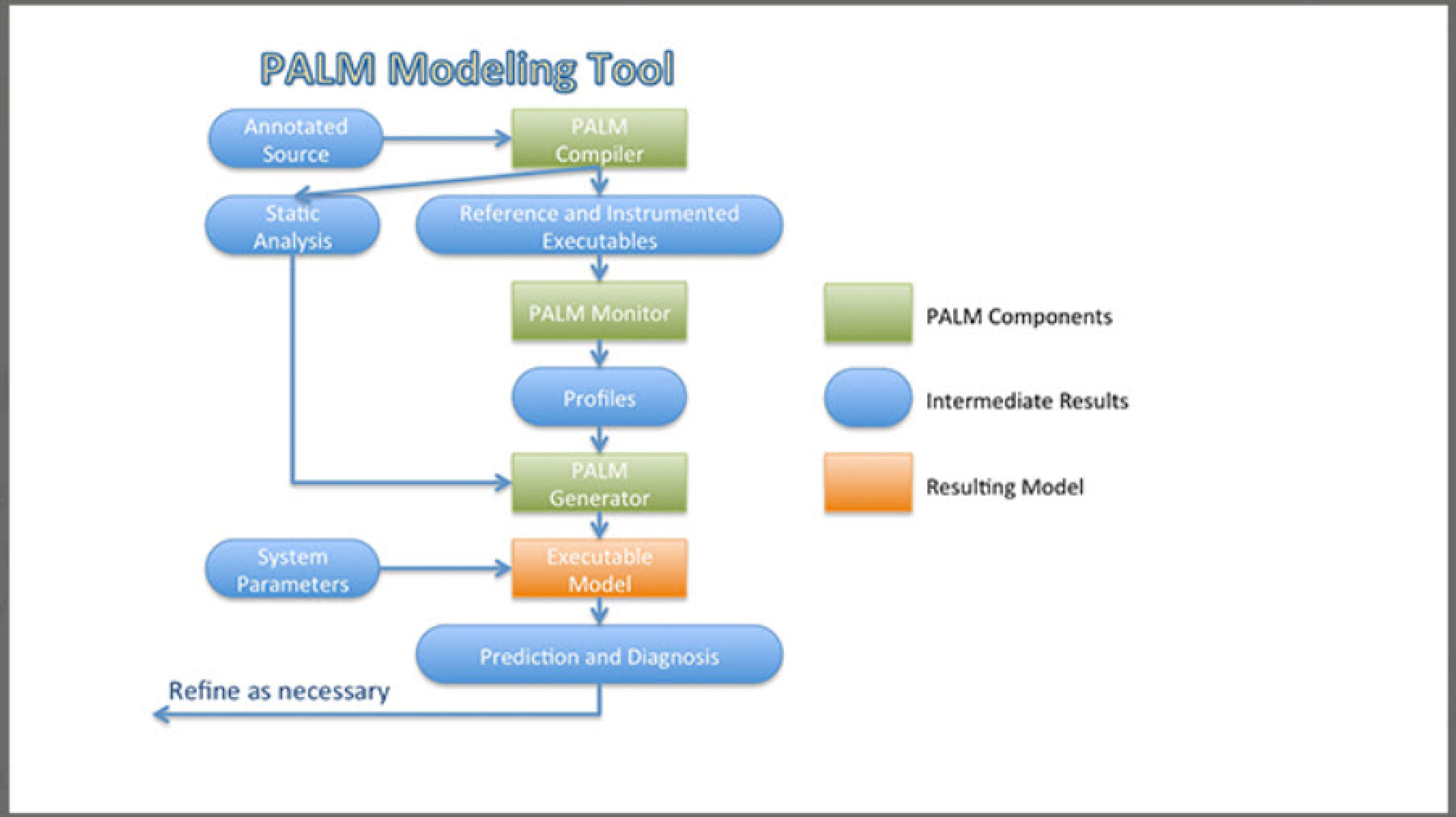
FUNDING & CREDITS

U.S. DEPARTMENT OF ENERGY

EX | COLL | ME

Paci





Behavio

Systems a

scale worklo
allow researc
n options tha

ng interesting

Behavioral analysis and mapping: characterizing key application behaviors (e.g., computation, memory access, communication) and mapping tasks onto resources.

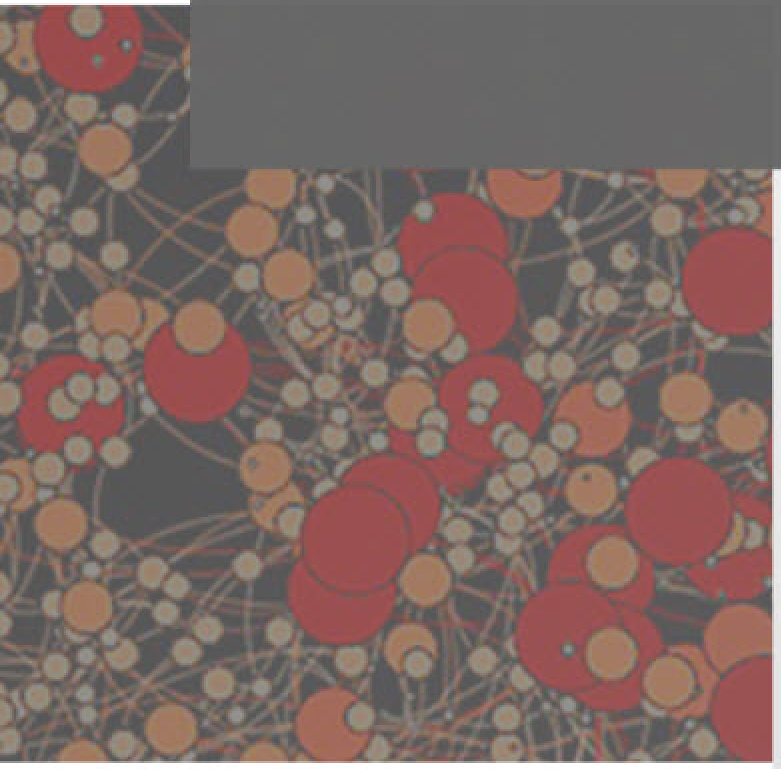
Prediction and validation: evaluation of models to derive predictions and validation of predictions against available measurements or simulations

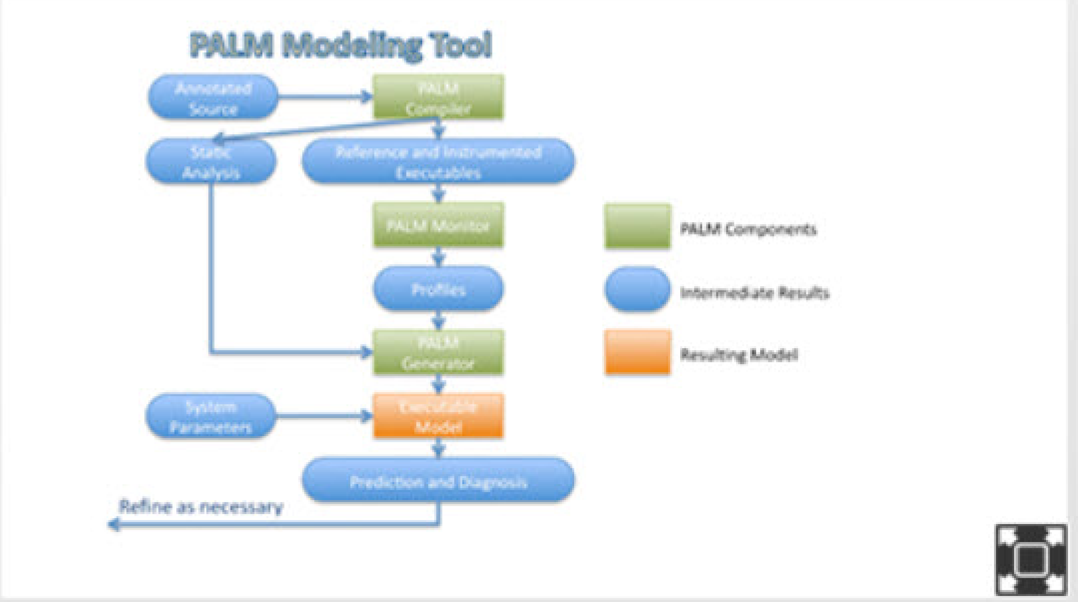
PNNL currently is advancing th o
distributed workflows, evalua of
Energy and Department of De

This work is foundational and has been funded by the U.S. Department of Energy's Office of Advanced Scientific Computing Research and the U.S. Department of Defense under various projects. PNNL is operated by Battelle for DOE under Contract DE-AC05-76RL01830. (PNNL-SA-114042)

Contact: Kevin J. Barker, kevin.barker@pnnl.gov

Modeling requires expertise a M tool aims to simplify the modeling e model and application can be develo models are executable, reproducible, and distributable.

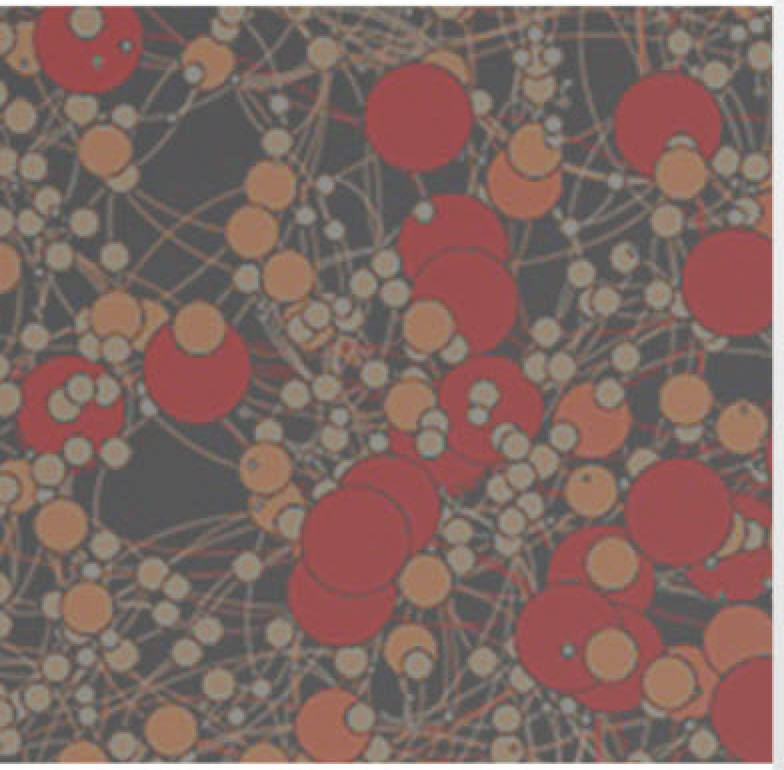




BACK TO MAP

FUNDING & CREDITS

Pacific Northwest NATIONAL LABORATORY



BACK TO MAP

Assessing the Impact of Silicon Photonics for Graph Applications

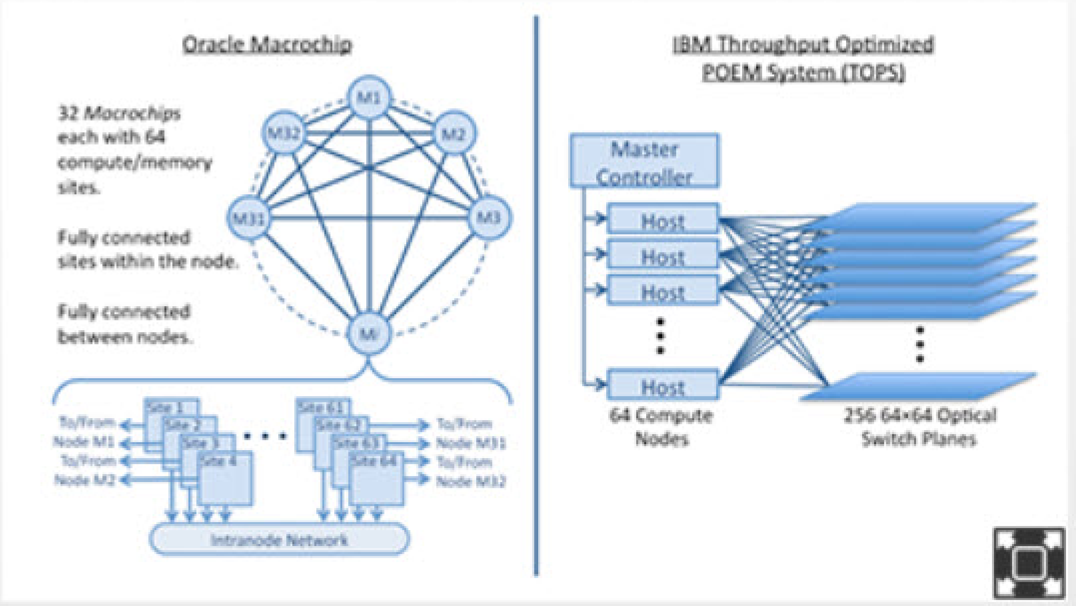
Using Performance Modeling to Evaluate Novel Technologies

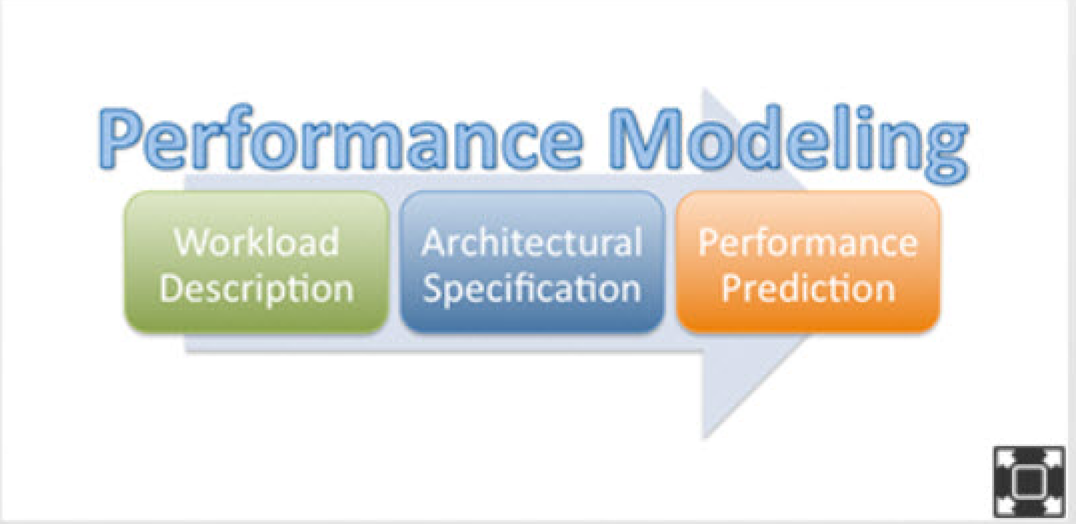
A key application of PNNL's performance modeling capability is in assessing the potential performance impact of future technologies. Silicon Photonics is on the near-term horizon and promises to deliver improved data transfer performance and energy efficiency over today's networking fabrics. PNNL has applied its performance modeling methodology to the question of whether or not these performance benefits can be realized for graph analytics applications.

Workload. PNNL's expertise yields representative examples of graph applications: Community Detection and Half-approximate Weighted Matching.

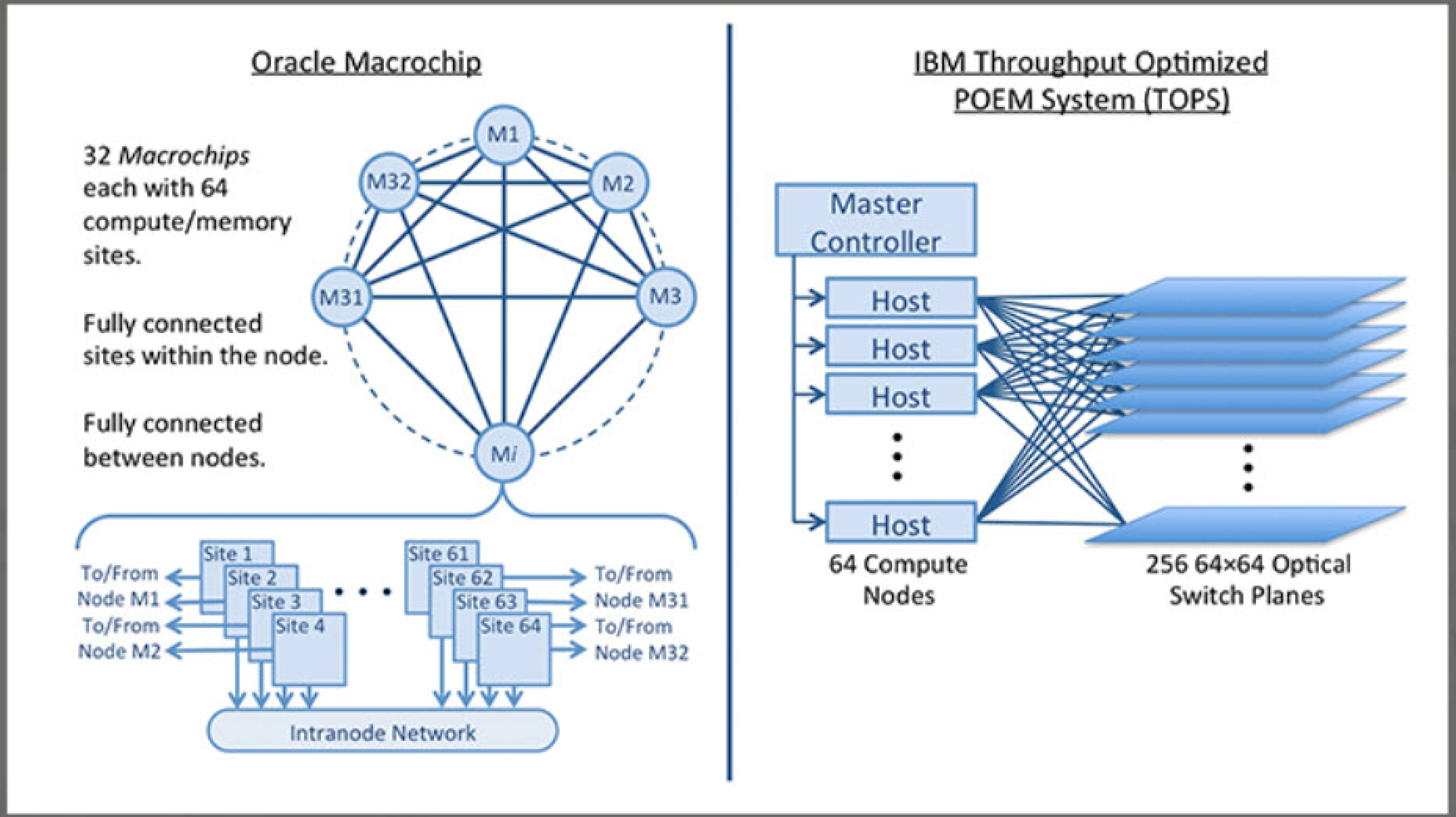
Architectures. The Oracle Macrochip employs a two-level optical fully connected network. IBM's TOPS system uses crossbar optical switch planes.

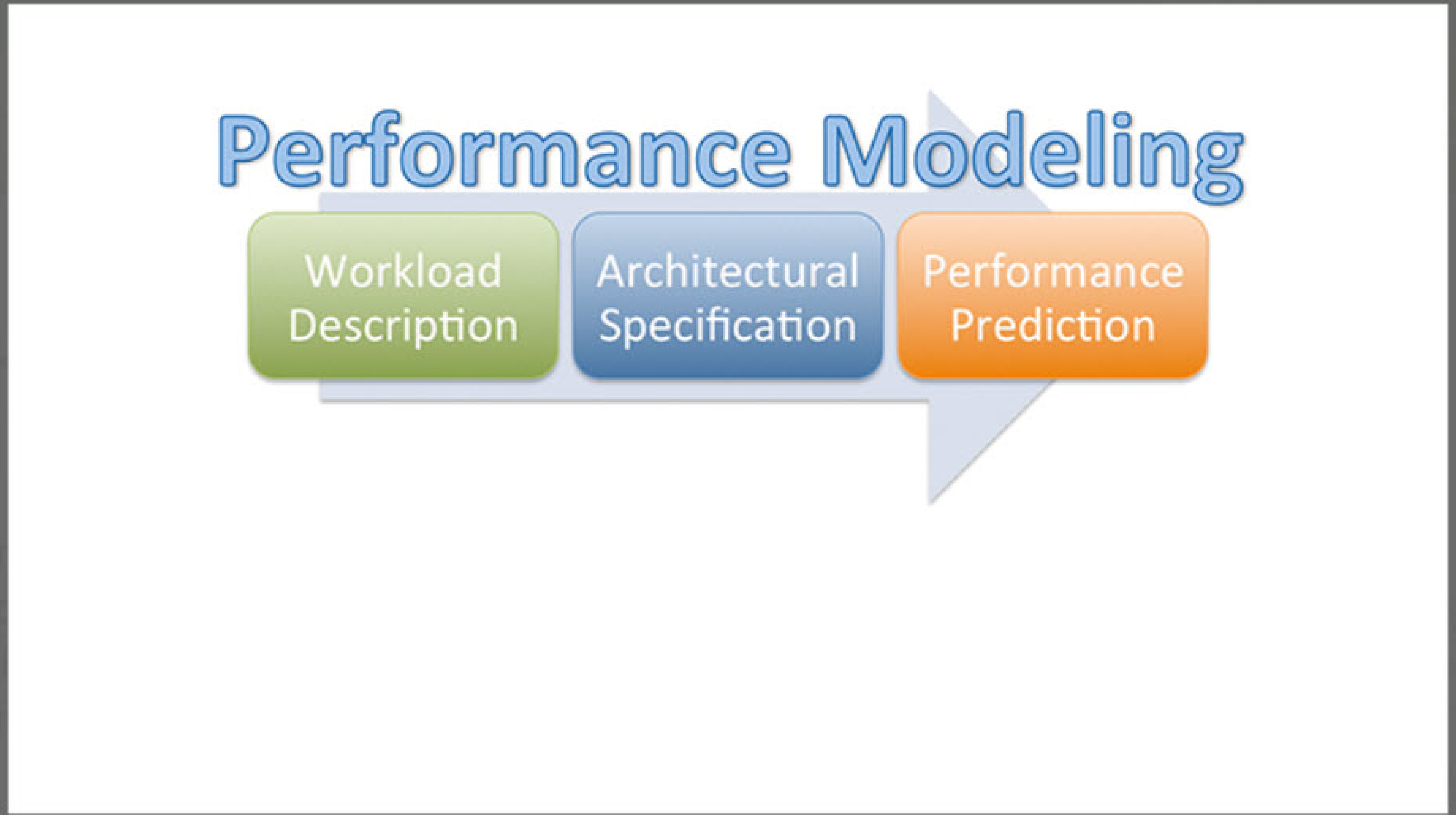
Modeling. Analytical modeling enables exploration of multiple design points in rapid succession. Workload-specific models capture key interactions between applications and architectures. By comparing and contrasting multiple designs, models can identify critical features or performance bottlenecks.





FUNDING & CREDITS





Paci

licon P
uate Nove

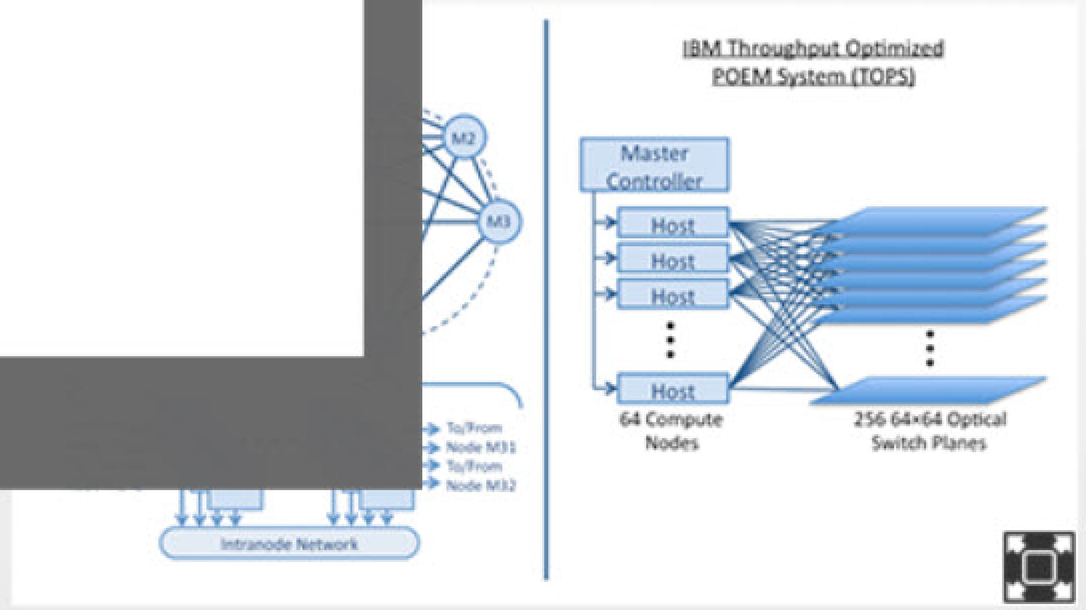
capability is
n horizon and
tworking fal
ese perform

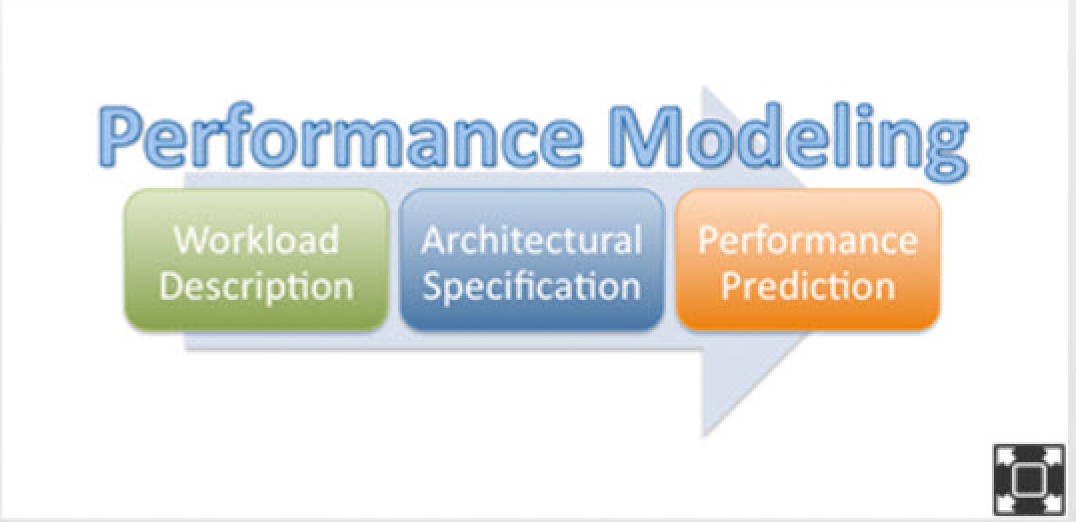
Workload. PNNL's expertise yields representative examples of graph applications: Community Detection and Half-approximate Weighted Matching.

Architectures. The Oracle Macrochip employs a two-level optical fully connected network. IBM's TOPS system uses crossbar optical switch planes.

Modeling. Analytical modeli enables exploration of multiple design points in rapid succession. Worklo ecific models capture key interacti ns between applications and architectures. By comparing and contrasting le designs, models can identify critical fe

This work was funded by the Defense Advanced Research Projects Agency under the Photonically Optimized Embedded Microprocessors (POEM) program. PNNL is operated by Battelle for DOE under Contract DE-AC05-76RL01830.
(PNNL-SA-114041)
Contact: Kevin J. Barker, kevin.barker@pnnl.gov

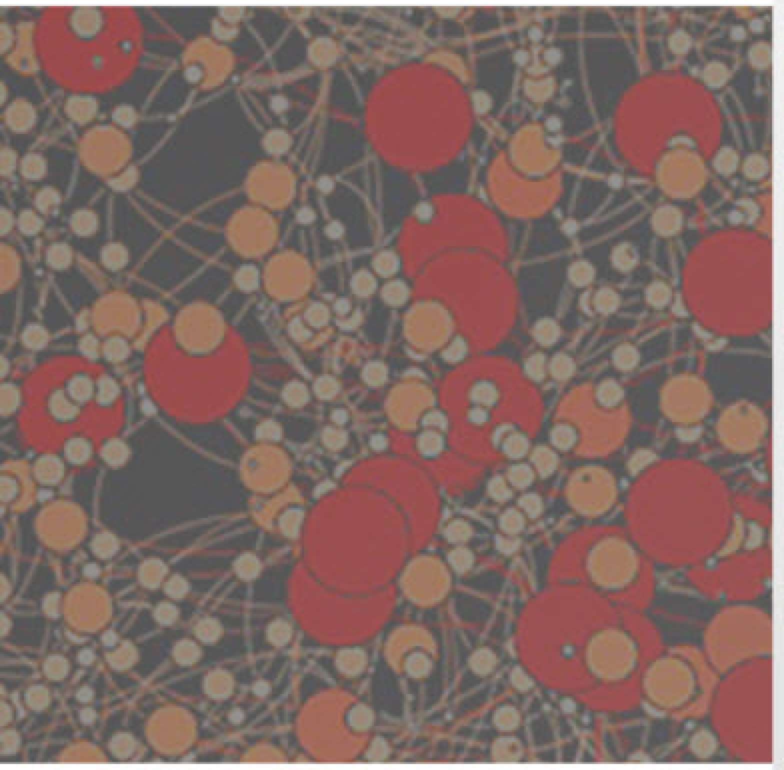




BACK TO MAP

FUNDING & CREDITS

Pacific Northwest NATIONAL LABORATORY



BACK TO MAP

High Performance Data Analytics

PNNL's Multifaceted Program to Accelerate Big Data Analytics using High Performance Computing

Introduced in 2013, the High Performance Data Analytics program, known as HPDA and led by Pacific Northwest National Laboratory, has been exploring, evaluating, and demonstrating the application of high-performance computing technologies to data analytics challenges.

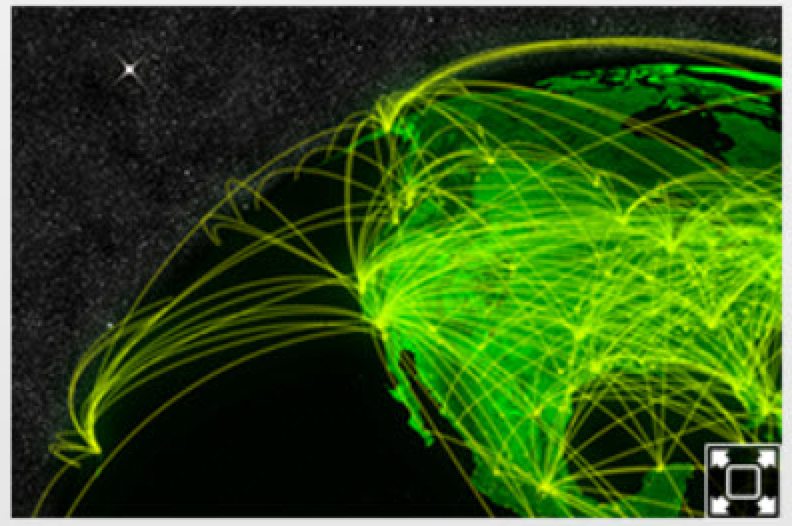
HPDA's core focus areas include:

Streaming Analytics: New algorithms and approaches to rapidly analyze high-bandwidth streaming data

Graph Analytics: Graph modeling, visualization, and evaluation for understanding large, complex networks

Compute Intensive Analytics: Novel hardware architectures and innovative techniques to solve computationally intensive analysis problems

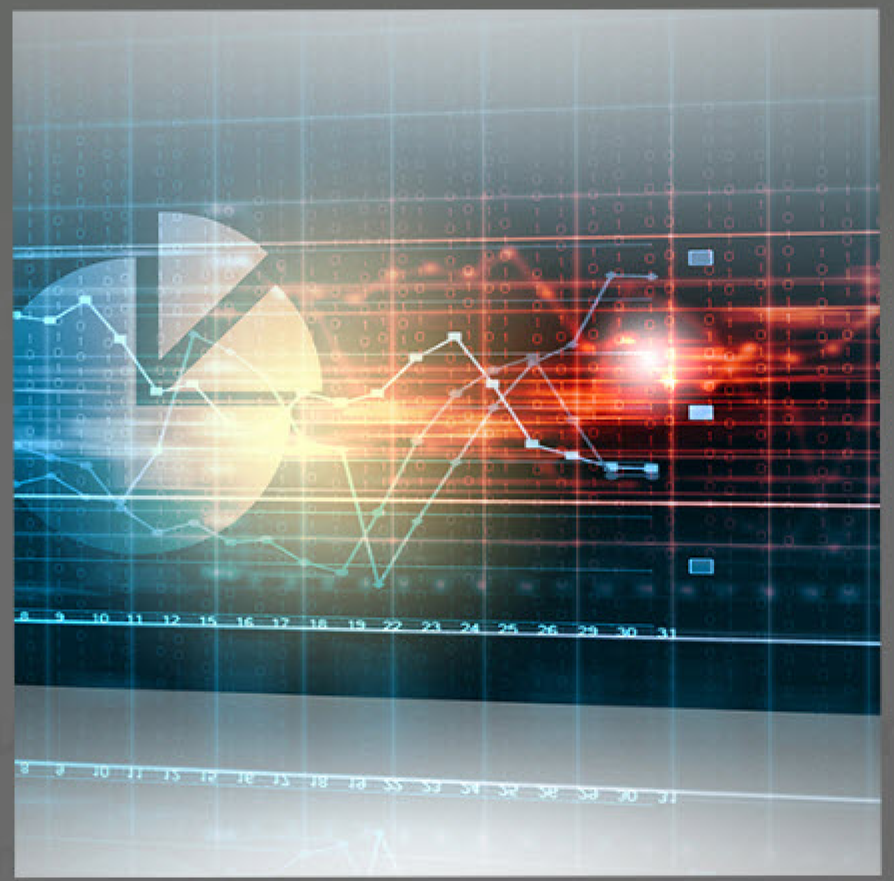
Exploratory Data Analysis: Mechanisms to explore and analyze massive, high-velocity data sources to gain new insights and inform decisions



Distributed Heterogeneous HPC Testbed: Evaluation, characterization, and modeling of emerging architectures

HPDA Partner Institutions

FUNDING & CREDITS



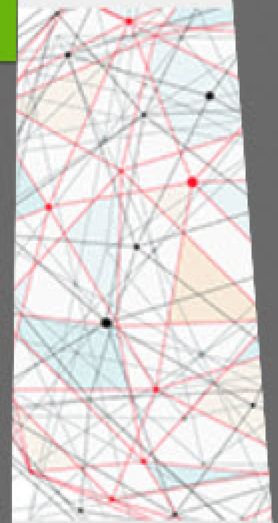
formance

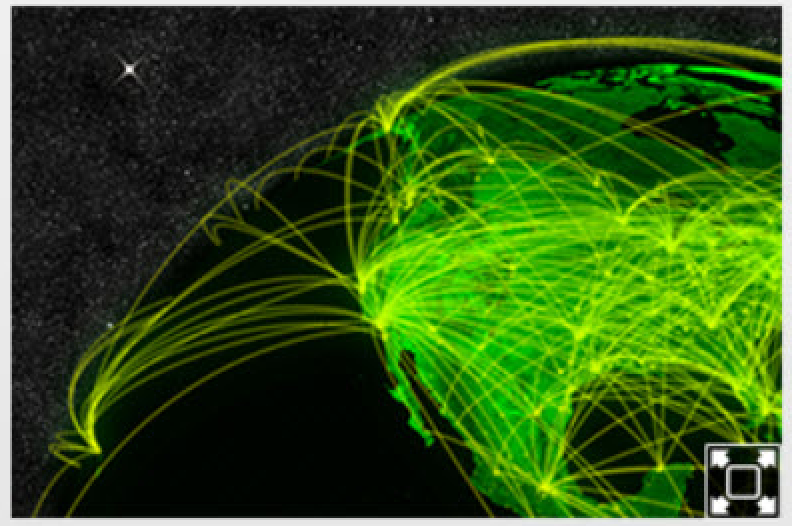
aceted Prog a Analytics mputing

Pacific Northwest
NATIONAL LABORATORY

Introduced in 2013, the Hig Performance Data Analytics program, known as HPDA tional Laboratory, has been exploring, evaluating, and demonstrating the application of high-performance puting technologies to data analytics challenges.

HPDA's core focus areas include:





ph Analytics: h modeling, v aluation for standing larg complex rks

solve computationally intensive analysis problems

ory Data Analysis: Mechanisms to explore and analyze massive, high-velocity data sources to gain new insights and inform decisions

Distributed Heterogeneous HPC Testbed: Evaluation, characterization, and modeling of emerging architectures

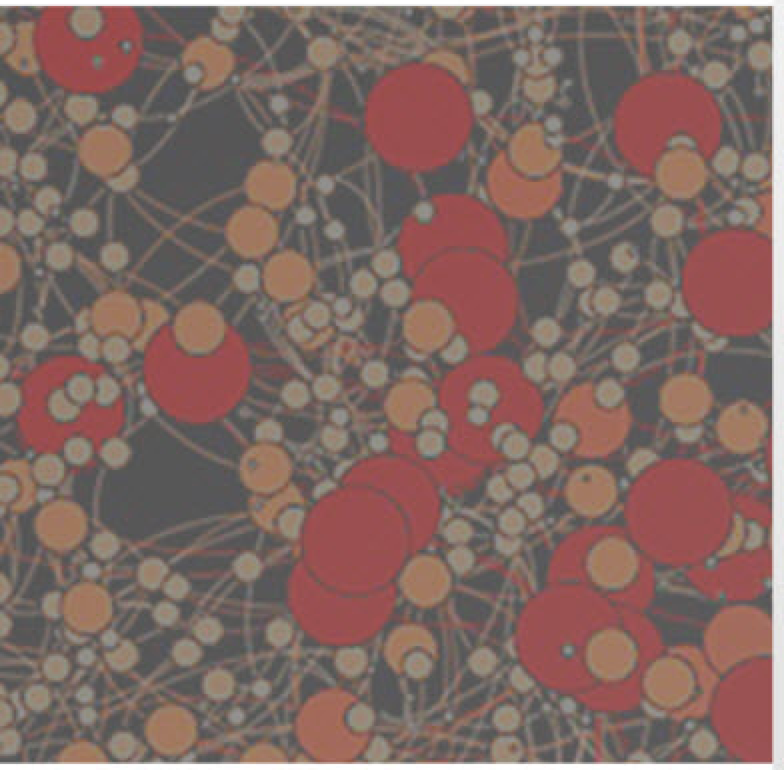
BACK TO MAP

HPDA Partner Institutions

High Performance Data Analytics Program is a collaboration led by PNNL with partners Mississippi State University, University of Washington, and Georgia Institute of Technology. PNNL is operated by Battelle for the U.S. Department of Energy under Contract DE-AC05-76RL01830. (PNNL-SA-114179)
Contact: John R. Johnson, john.johnson@pnnl.gov

FUNDING & CREDITS

Pacific Northwest NATIONAL LABORATORY



BACK TO MAP

Integrated End-to-End Performance Prediction and Diagnosis for Extreme Scientific Workflows (IPPD)

Using Performance Modeling to Evaluate Novel Technologies

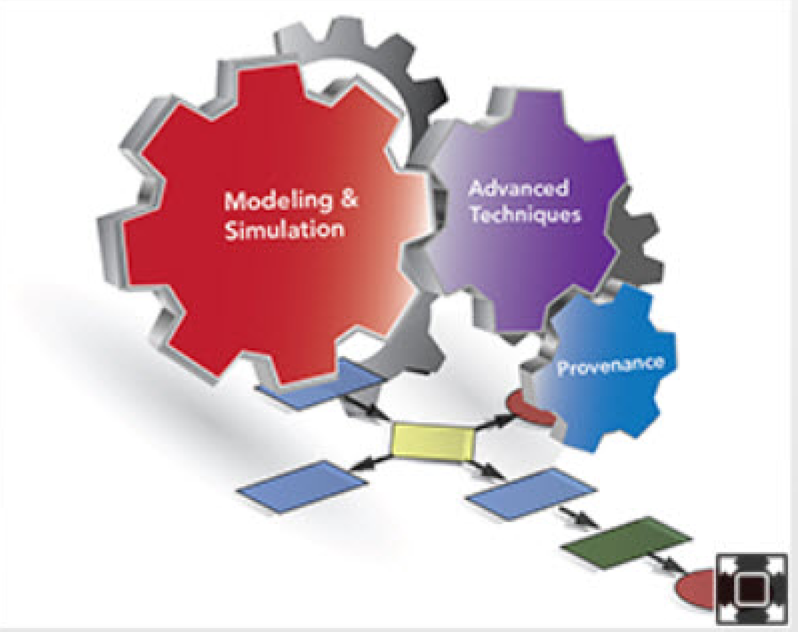
Darren J Kerbyson (PI), Ilkay Altintas[1], Kevin Barker, Jeff Daily, Ryan Friese, Mahantesh Halappanavar, Kerstin Kleese van Dam[2], Mai Nguyen, Malachi Schram, Alok Singh[1], Eric Stephan, Nathan Tallent, and Jian Yin
([1]University of California, San Diego Supercomputing Center; [2]Brookhaven National Laboratory)

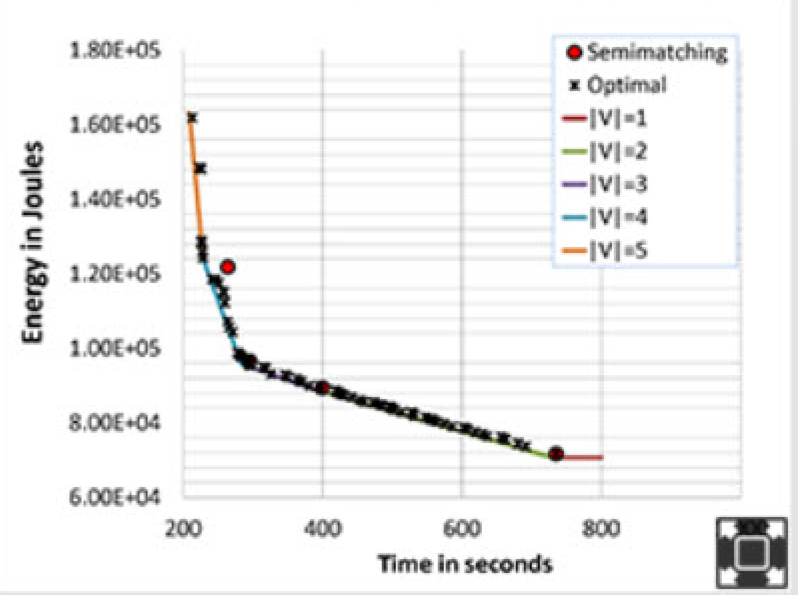
IPPD aims to provide an integrated approach to the modeling and execution of extreme-scale scientific workflows. IPPD brings together researchers and domain scientists working on modeling and simulation, empirical analysis, optimization, workflows. IPPD builds on existing research focused on high-performance computing systems and applications and currently is using workflows from high-energy physics and climate modeling as case studies.

Modeling and Simulation. Develop modeling and simulation to enable both exploration in advance of possible workflow configuration and optimizations, as well as rapid performance prediction to guide dynamic adaptation of workflows and optimization of resource utilization. We generate models from static/dynamic annotation structure using the PALM modeling tool and use black-box modeling approaches.

Provenance. Capture empirical performance information from workflows enabling baseline performance to be established, as well as to identify and diagnose variability and feed simulation, modeling, and optimization. To enable in-depth performance analysis, we extend the Open Provenance Model to capture performance information for workflows and model runs.

Advanced Techniques. Explore novel techniques for workflow optimization for both processing and data organization. We build on the unit-commitment analogy in power grids to select a cost-efficient set of resources that meet the forecasted demand and develop Pareto-optimal solutions for energy-efficient resource use. For optimal data movement and storage, we also use transparent page management and compression.





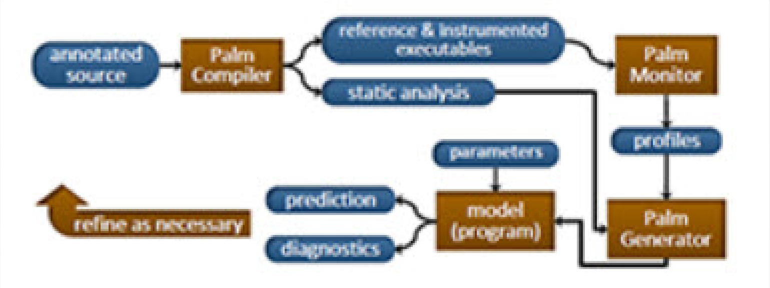


Figure 1: Workflow of the Palm modeling tool.

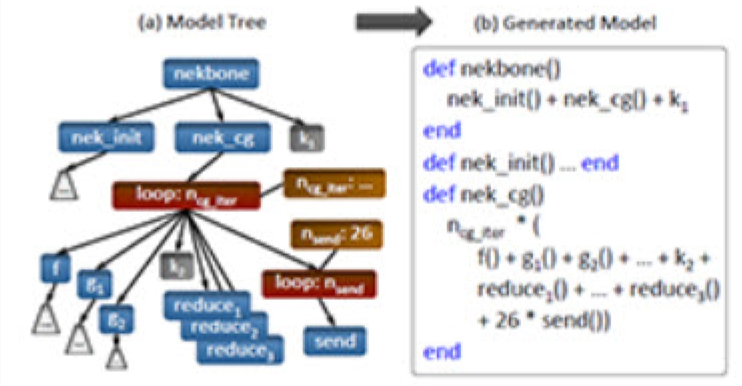


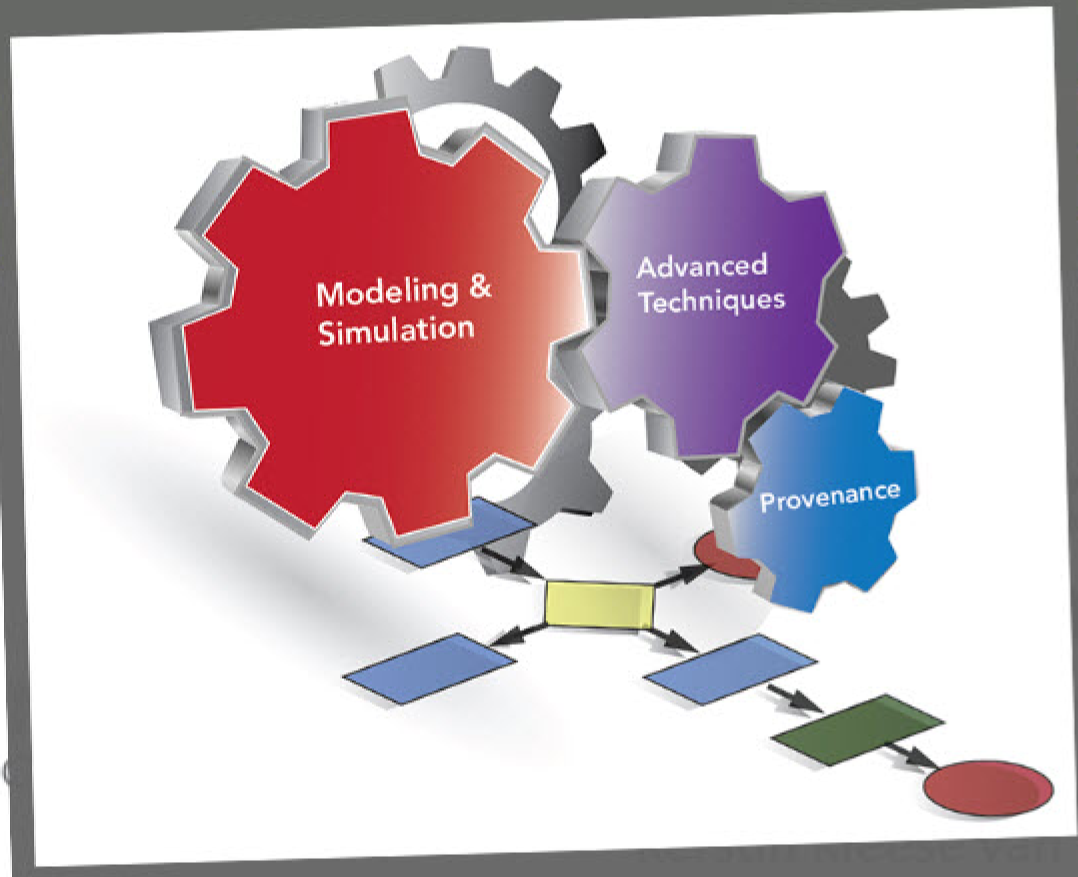
Figure 2: Palm's Generator uses (a) a Model Tr... to synthesize (b) a model (a program).

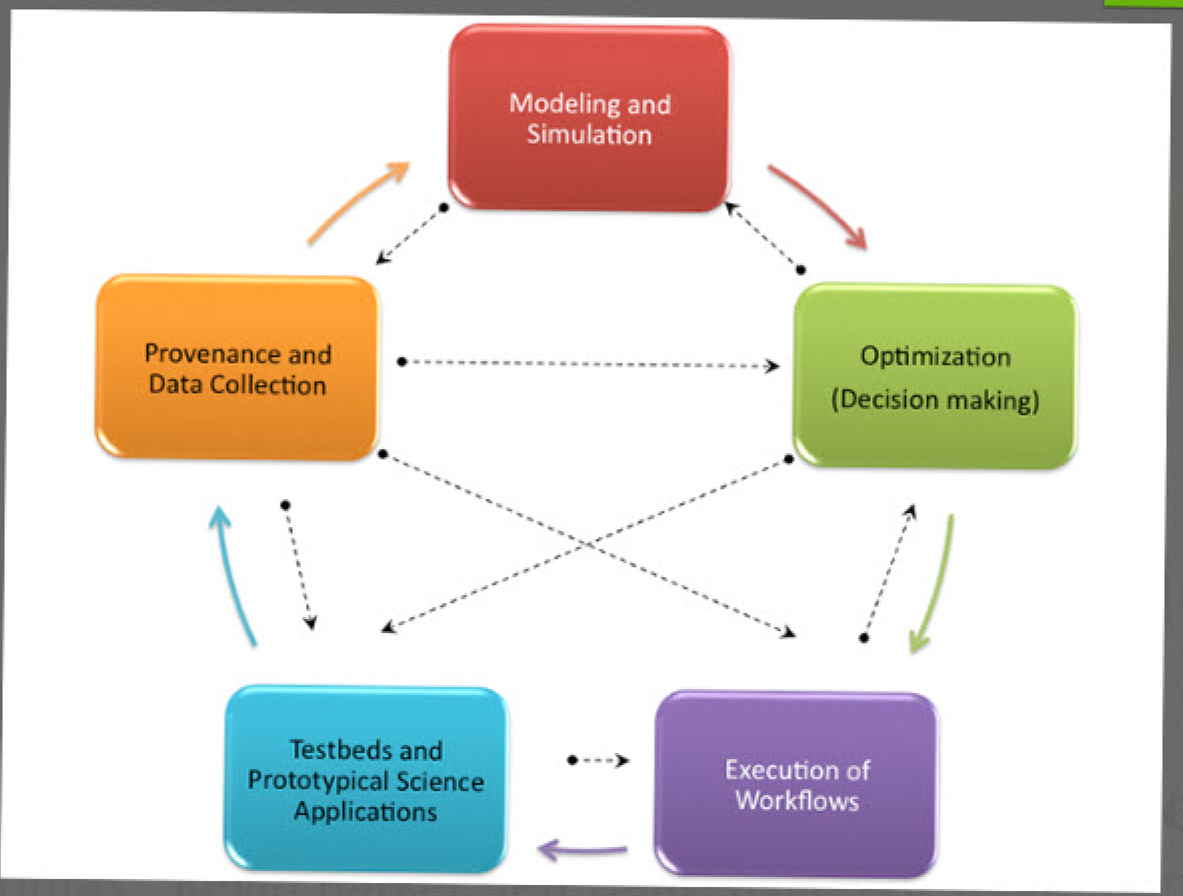
FUNDING & CREDITS

The National Laboratory System

DOE ... PC

EXPL... EXPLOR... PROJECTS | B...





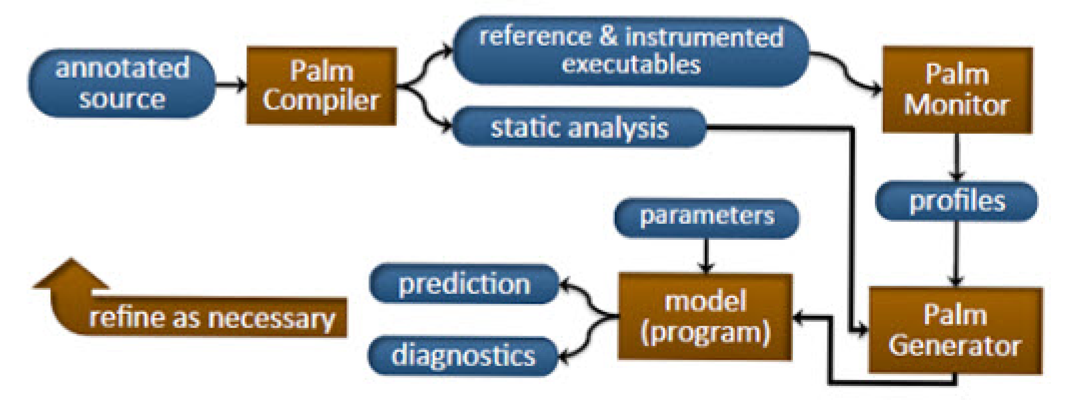


Figure 1: Workflow of the Palm modeling tool.

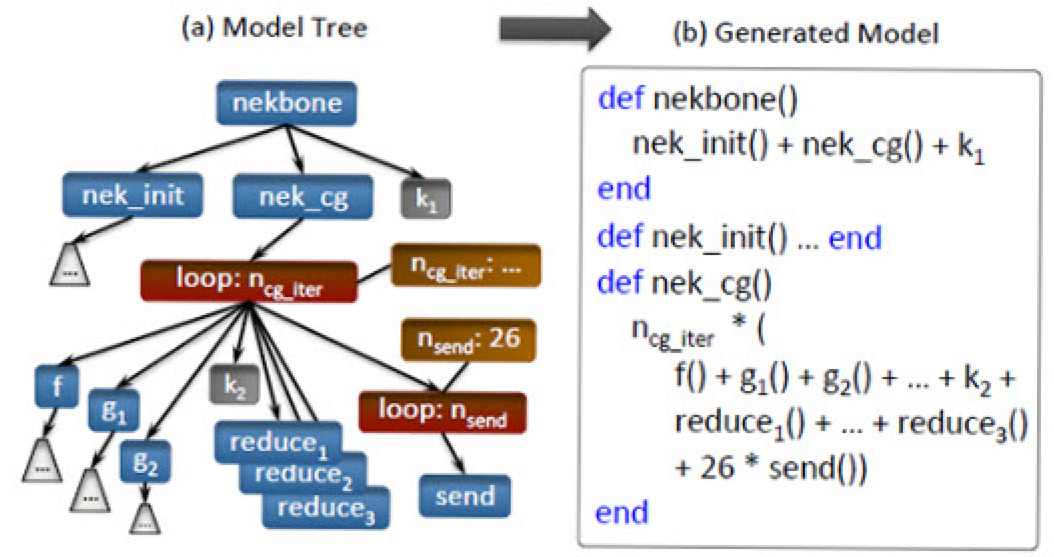


Figure 2: Palm's Generator uses (a) a Model Tree, to synthesize (b) a model (a program).

...l End-... ...ction and Dia...

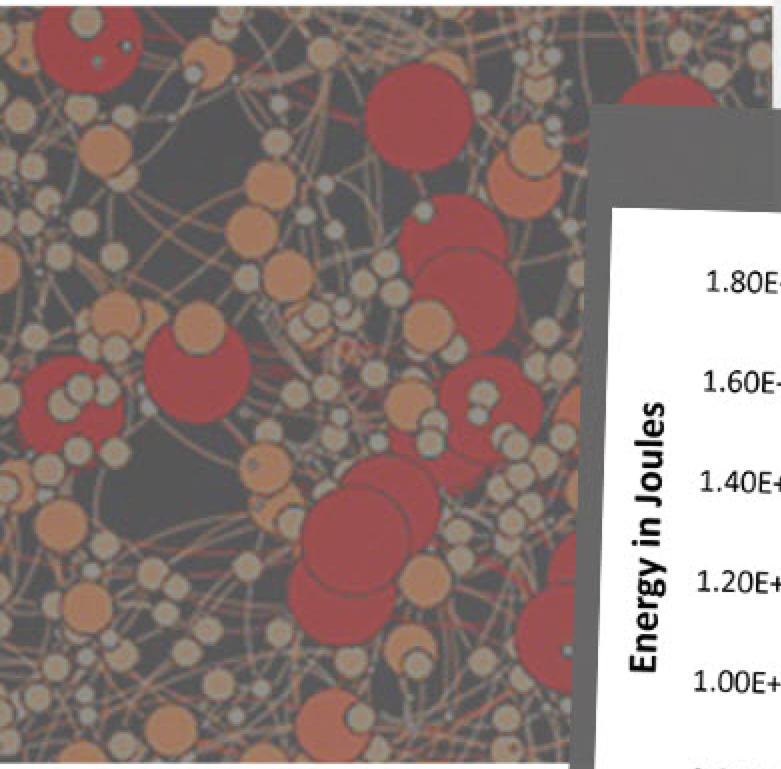
...e Sci...

...ance M... ...es

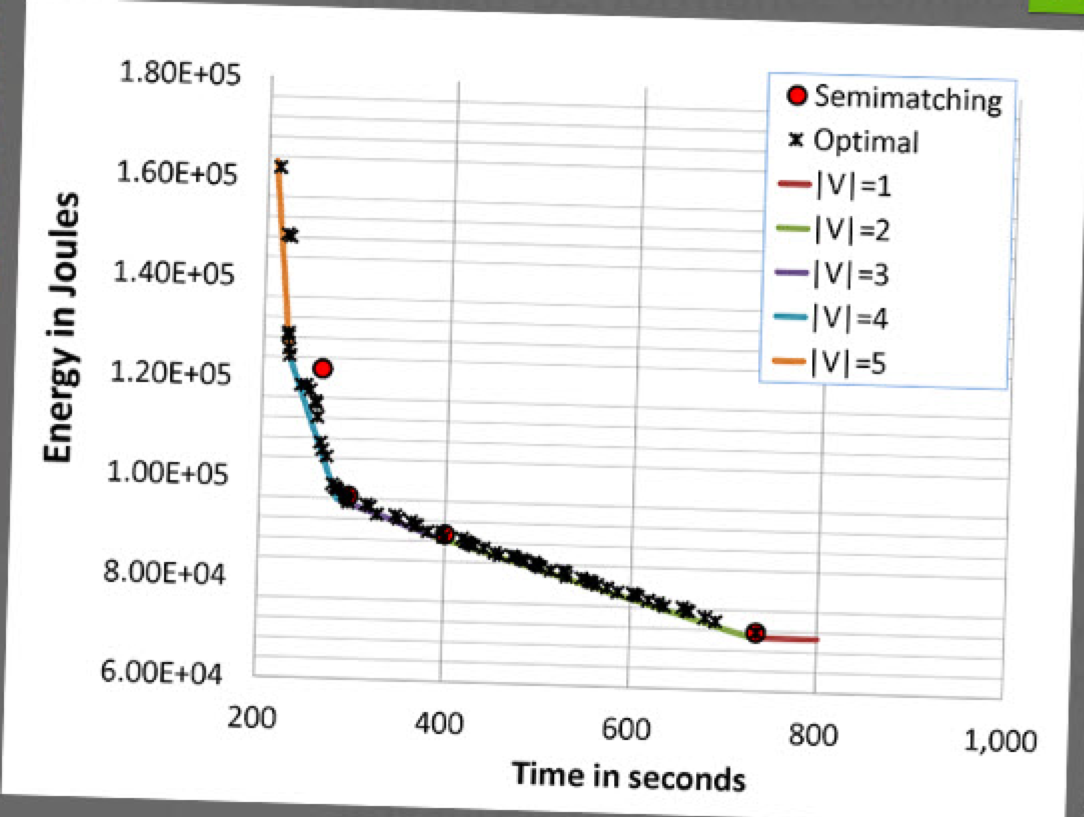
...PI), Ilkay A... ...tesh Halappanavar,
...am[2], Mai... ...Nathan Tallent, and Jian Yin
...San Diego Supercomputing Center; ...

Pacific

NATI...



IPPD aims to provide an integrated approach to the modeling and execution of extreme-scale scientific workflows. IPPD brings together researchers and domain scientists working on modeling and simulation, empirical analysis, optimization, workflows. IPPD builds on existing research focused on high-performance compu... systems and applications and currently is using workflows from ... modeling as case studies.

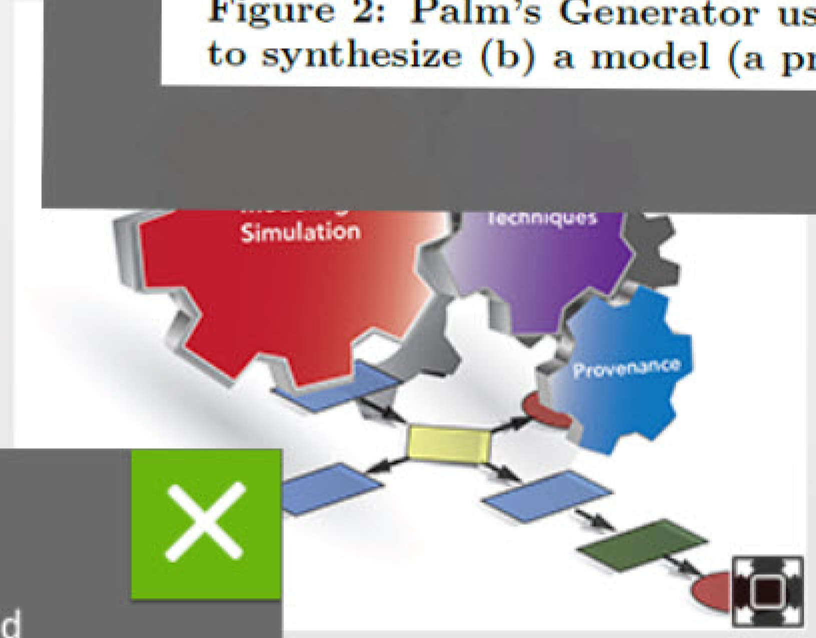


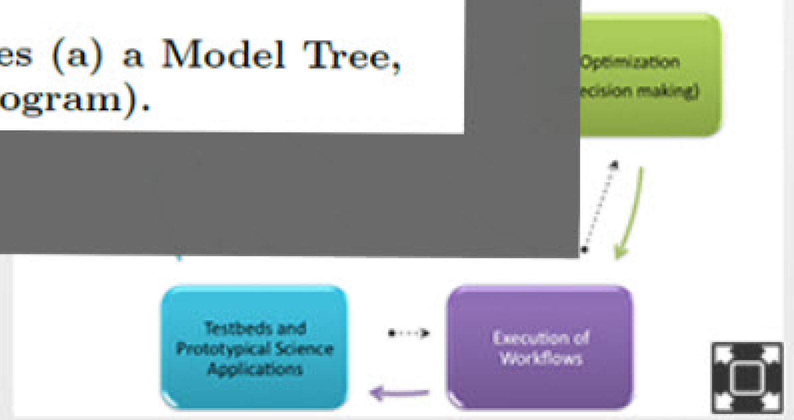
...modeling and simulation to enable both exploration in advance of possible ...ations, as well as rapid performance prediction to guide dynamic adaptation ...source utilization. We generate models from static/dynamic annotation ...tool and use blackbox modeling approaches.

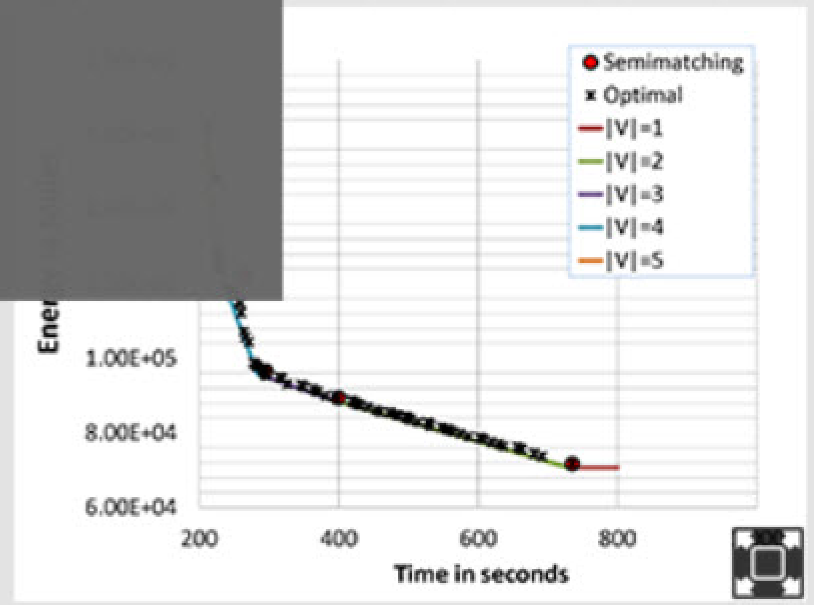
This work was funded by the U.S. Department of Energy Awards FWP-66406 and DE-SC0012630. PNNL is operated by Battelle for DOE under Contract DE-AC05-76RL01830. (PNNL-SA-114141)
Contact: Darren J. Kerbyson, darren.kerbyson@pnnl.gov

...el techniques for workflow optimization for both processing and data ... commitment analogy in power grids to select a cost-efficient set of resources that meet the forecasted demand and develop Pareto-optimal solutions for energy-efficient resource use. For optimal data movement and storage, we also use transparent page management and compression.

BACK TO MAP







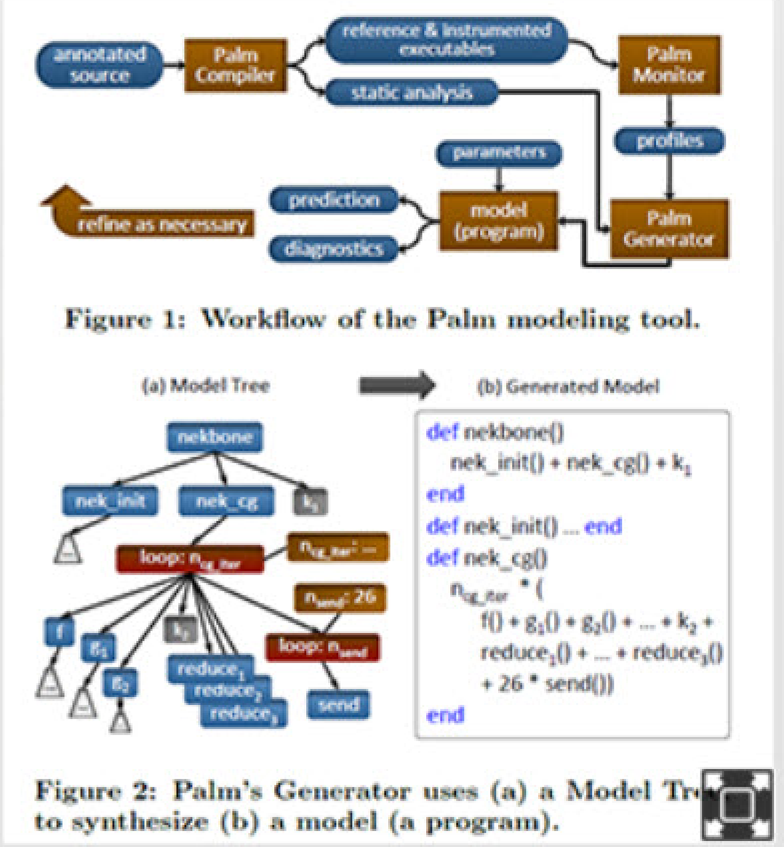


Figure 1: Workflow of the Palm modeling tool.

FUNDING & CREDITS